



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2018

---

## **Robust Utility Maximization in Discrete-Time Markets with Friction**

Neufeld, Ariel ; Sikic, Mario

**Abstract:** We study a robust stochastic optimization problem in the quasi-sure setting in discrete-time. We show that under a linearity-type condition the problem admits a maximizer. This condition is implied by the no-arbitrage condition in models of financial markets. As a corollary, we obtain existence of a utility maximizer in the frictionless market model, markets with proportional transaction costs and also more general convex costs, like in the case of market impact.

DOI: <https://doi.org/10.1137/16M1101829>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-185077>

Journal Article

Published Version

Originally published at:

Neufeld, Ariel; Sikic, Mario (2018). Robust Utility Maximization in Discrete-Time Markets with Friction. SIAM Journal on Control and Optimization, 56(3):1912-1937.

DOI: <https://doi.org/10.1137/16M1101829>

# ROBUST UTILITY MAXIMIZATION IN DISCRETE-TIME MARKETS WITH FRICTION\*

ARIEL NEUFELD<sup>†</sup> AND MARIO ŠIKIĆ<sup>‡</sup>

**Abstract.** We study a robust stochastic optimization problem in the quasi-sure setting in discrete-time. We show that under a linearity-type condition the problem admits a maximizer. This condition is implied by the no-arbitrage condition in models of financial markets. As a corollary, we obtain existence of a utility maximizer in the frictionless market model, markets with proportional transaction costs and also more general convex costs, like in the case of market impact.

**Key words.** robust optimization, utility maximization, financial markets with friction

**AMS subject classifications.** 93E20, 49L20, 91B28

**DOI.** 10.1137/16M1101829

**1. Introduction.** An agent participates in the market by buying and selling options. If we denote the portfolio by  $H$ , i.e., the holdings of the agent at time  $t$  by  $H_t$ , then after  $T$  steps and liquidation of the portfolio, he or she will have amount  $V(H)$  in the bank account. Instead of moving arbitrarily in the market, we ask ourselves whether the optimal portfolio process, or strategy, of the trader exists. When talking about optimality, we need to specify a preference relation on the set of final wealth or states of the bank account. The preference we will be considering is the robust utility preference

$$(1) \quad H = (H_0, \dots, H_{T-1}) \mapsto \inf_{P \in \mathfrak{P}} E^P [U(V(H_0, \dots, H_{T-1}))],$$

where the utility function is given by  $U$  and  $\mathfrak{P}$  denotes a collection of probability measures.

The use of robust utility preference is motivated by the fact that the true probability measure might not be known. So, instead of considering the utility maximization with respect to one measure, which we guess to be the correct one, we consider a family of probability measures that are possible. We then talk about model uncertainty: it is not known which of the probability measures is the true one, but hopefully, the class of probability measures we are considering is big enough to contain the true one. The optimal strategy for the robust utility maximization problem is giving “the best possible performance” under “the worst probability measure.”

Instead of specifying the market model and analyzing the robust utility maximization problem, it will be easier to consider a more general model. The advantage of doing this, besides simpler notation, is the fact that not every utility optimization problem is of the type: maximizing the robust utility of terminal wealth. One can, for instance, consider the problem of optimal consumption stream. We will, thus, consider the following robust optimization problem:

$$(2) \quad \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P [\Psi(H_0, \dots, H_{T-1})].$$

\*Received by the editors November 2, 2016; accepted for publication (in revised form) January 19, 2018; published electronically May 31, 2018.

<http://www.siam.org/journals/sicon/56-3/M110182.html>

**Funding:** This work was funded by Swiss National Foundation grant SNF 200020\_172815.

<sup>†</sup>Department of Mathematics, ETH Zurich, Zurich, Switzerland (ariel.neufeld@math.ethz.ch).

<sup>‡</sup>Center for Finance and Insurance, University of Zurich, Zurich, Switzerland (mario.sikic@bf.uzh.ch).

The goal of this paper is to show the existence of a maximizer  $\hat{H}$  for a general class of concave functions  $\Psi$ . In section 3, we will provide examples showing that this more general problem indeed contains the robust utility maximization problem. The examples include the liquidation value of the strategy  $H$  at maturity  $T$  in frictionless market models, but also with proportional transaction costs and even more general costs.

Concerning the set  $\mathfrak{P}$ , there are two cases that prompt different approaches. If the set  $\mathfrak{P}$  is dominated, i.e., there exists a probability measure  $P$  such that every measure  $Q \in \mathfrak{P}$  is absolutely continuous with respect to  $P$ , there is a plethora of approaches to obtain existence of optimizers. A simple one, which works under our set of assumption is to start with the maximizing sequence of strategies and pass, using a Komlos-type lemma (see, e.g., [8, Lemma 9.8.1, p. 202]), to a convergent sequence of convex combinations. The limit of this sequence, thus, needs to be the maximizer. One can also apply dynamic programming, like in [22], to obtain existence under a relaxed set of assumptions, or a duality argument; see [19].

If the set  $\mathfrak{P}$  of probability measures on  $\Omega$  is not dominated, the set of approaches to establish existence becomes very limited. Many approaches of the dominated case do not transfer over to this one; for instance, there is no analogue of a Komlos-type lemma. So, the simple argument for existence provided above does not work anymore. Also, the duality approach is, to the best of our knowledge, not applied here. What remains is the dynamic programming approach. For that reason, the setup we are working in needs to be carefully laid.

Robust utility maximization was already considered in the literature. The closest to our work in discrete-time is [18], which shows existence of an optimal strategy in the frictionless market model where the utility function is defined on the positive half-line. This work is essentially the robust analogue of [22]. For more results concerning the robust utility maximization problem in a nondominated framework, we refer to [2, 4, 7, 9, 12, 15, 16, 17, 26]. To the best of our knowledge, robust utility maximization in the nondominated setting for financial markets with friction has not been studied yet.

The main tool for proving existence is dynamic programming. Dynamic programming is an approach that replaces the multistep decision problem with a series of one-step decision problems. If one can solve, i.e., prove existence of optimizers of the one-step problems, then one gets existence in general, by using those one-step optimizers in sequence. In order to be able to apply the approach, we need to first work in the setup that is suitable for measurable selection. The formulation of a robust market from [6] turns out to be appropriate for our needs. We will follow the formulation of dynamic programming in [11].

How does one establish the existence of the optimal strategy in the one-step models that we need to solve? One approach, taken in [18] and [22], is to set up the problem in such a way that the set of strategies in the one-step problems one obtains is compact. Indeed, under a suitable no-arbitrage condition, having a frictionless market model with utility function defined on the positive half-line implies this compactness, up to the projection on the predictable range.

To consider utility functions that are defined on the whole real line, for example, the exponential utility function, one needs to resort to convex analysis. Theorem 9.2 in [23] provides the necessary condition. When considering utility maximization in a one-step market model, this condition turns out to be just the no-arbitrage condition; we will prove this statement in the course of the paper. In our general setting, the

condition that will serve as a no-arbitrage condition will be the following:

$$(3) \quad \text{the set } \mathcal{K} := \{H \in \mathcal{H} \mid \Psi^\infty(H_0, \dots, H_{T-1}) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\} \quad \text{is linear,}$$

where  $\Psi^\infty$  denotes the horizon function of the concave function  $\Psi$ ; see section 2. This condition coincides with the robust no-arbitrage condition in the frictionless market model. The main step, indeed the main technical obstacle in the proof of our main result, is to show that the “global” condition (3) satisfies a “local” version at each time step.

The remainder of this paper is organized as follows. In section 2, we introduce the concepts, list the assumptions imposed on  $\Psi$ , and state the main results. The examples of robust utility maximization in different financial markets with friction are given in section 3. In section 4, we introduce and solve the corresponding one-period maximization problem. The problem of finding a maximizer reduces to a question of closedness property of the hypograph of the function over which one maximizes. The result [23, Theorem 9.2, p. 75] provides an answer to the above question and explains the sufficiency of the condition (3). In section 5, we introduce the notion needed in our dynamic programming approach and explain why this leads to the existence of a maximizer in our optimization problem (2). The proof is then divided into several steps, which heavily use the theory of lower semianalytic functions.

*Notation.* For any vector  $x \in \mathbb{R}^{dT}$ , written out as  $x = (x_0, \dots, x_{T-1})$ , where  $x_i \in \mathbb{R}^d$  for each  $i$ , we denote the restriction to the first  $t$  entries by  $x^t := (x_0, \dots, x_{t-1})$ . For  $y \in \mathbb{R}^{dT}$ , we denote by  $x \cdot y$  the usual scalar product on  $\mathbb{R}^{dT}$ .

**2. Optimization problem.** Let  $T \in \mathbb{N}$  denote the fixed finite time horizon and let  $\Omega_1$  be a Polish space. Denote by  $\Omega^t := \Omega_1^t$  the  $t$ -fold Cartesian product for  $t = 0, 1, \dots, T$ , where we use the convention that  $\Omega^0$  is a singleton. Let  $\mathcal{F}_t := \bigcap_P \mathcal{B}(\Omega^t)^P$  be the universal completion of the Borel  $\sigma$ -field  $\mathcal{B}(\Omega^t)$ ; here  $\mathcal{B}(\Omega^t)^P$  denotes the  $P$ -completion of  $\mathcal{B}(\Omega^t)$  and  $P$  ranges over the set  $\mathfrak{M}_1(\Omega^t)$  of all probability measures on  $(\Omega^t, \mathcal{B}(\Omega^t))$ . Moreover, define  $(\Omega, \mathcal{F}) := (\Omega^T, \mathcal{F}_T)$ . This plays the role of our initial measurable space.

For every  $t \in \{0, 1, \dots, T-1\}$  and  $\omega^t \in \Omega^t$  we fix a nonempty set  $\mathfrak{P}_t(\omega^t) \subseteq \mathfrak{M}_1(\Omega_1)$  of probability measures;  $\mathfrak{P}_t(\omega^t)$  represents the possible laws for the  $t$ th period given state  $\omega^t$ . Endowing  $\mathfrak{M}_1(\Omega_1)$  with the usual topology induced by the weak convergence makes it into a Polish space; see [3, Chapter 7]. We assume that for each  $t$

$\text{graph}(\mathfrak{P}_t) := \{(\omega^t, P) \mid \omega^t \in \Omega^t, P \in \mathfrak{P}_t(\omega^t)\}$  is an analytic subset of  $\Omega^t \times \mathfrak{M}_1(\Omega_1)$ .

Recall that a subset of a Polish space is called analytic if it is the image of a Borel subset of a (possibly different) Polish space under a Borel measurable mapping (see [3, Chapter 7]); in particular, the above assumption is satisfied if  $\text{graph}(\mathfrak{P}_t)$  is Borel. The set  $\text{graph}(\mathfrak{P}_t)$  being analytic provides the existence of an universally measurable kernel  $P_t : \Omega^t \rightarrow \mathfrak{M}_1(\Omega_1)$  such that  $P_t(\omega^t) \in \mathfrak{P}_t(\omega^t)$  for all  $\omega^t \in \Omega^t$  by the Jankov-von Neumann theorem; see [3, Proposition 7.49, p. 182]. Given such a kernel  $P_t$  for each  $t \in \{0, 1, \dots, T-1\}$ , we can define a probability measure  $P$  on  $\Omega$  by

$$P(A) := \int_{\Omega_1} \cdots \int_{\Omega_1} \mathbf{1}_A(\omega_1, \dots, \omega_T) P_{T-1}(\omega_1, \dots, \omega_{T-1}; d\omega_T) \cdots P_0(d\omega_1), \quad A \in \mathcal{F},$$

where we write  $\omega := \omega^T := (\omega_1, \dots, \omega_T)$  for any element in  $\Omega$ . We denote a probability measure defined as above by  $P = P_0 \otimes \cdots \otimes P_{T-1}$ . For the multiperiod market, we

consider the set

$$\mathfrak{P} := \{P_0 \otimes \cdots \otimes P_{T-1} \mid P_t(\cdot) \in \mathfrak{P}_t(\cdot), t = 0, \dots, T-1\} \subseteq \mathfrak{M}_1(\Omega)$$

of probability measures representing the uncertainty of the law, where in the above definition each  $P_t : \Omega^t \rightarrow \mathfrak{M}_1(\Omega_1)$  is universally measurable such that  $P_t(\omega^t) \in \mathfrak{P}_t(\omega^t)$  for all  $\omega^t \in \Omega^t$ .

We will often interpret  $(\Omega^t, \mathcal{F}_t)$  as a subspace of  $(\Omega, \mathcal{F})$  in the following way. Any set  $A \subset \Omega^t$  can be extended to a subset of  $\Omega^T$  by adding  $(T-t)$  products of  $\Omega_1$ , i.e.,  $A^T := A \times \Omega_1 \times \cdots \times \Omega_1 \subset \Omega^T$ . Then, for every measure  $P = P_0 \otimes \cdots \otimes P_{T-1} \in \mathfrak{P}$ , one can associate a measure  $P^t$  on  $(\Omega^t, \mathcal{F}^t)$  such that  $P^t[A] = P[A^T]$  by setting  $P^t := P_0 \otimes \cdots \otimes P_{t-1}$ .

We call a set  $A \subseteq \Omega$   $\mathfrak{P}$ -polar if  $A \subseteq A'$  for some  $A' \in \mathcal{F}$  such that  $P[A'] = 0$  for all  $P \in \mathfrak{P}$ , and we say a property holds  $\mathfrak{P}$ -quasi surely, or simply  $\mathfrak{P}$ -q.s., if the property holds outside a  $\mathfrak{P}$ -polar set.

A map  $\Psi : \Omega \times \mathbb{R}^{dT} \rightarrow \overline{\mathbb{R}}$  is called an  $\mathcal{F}$ -measurable normal integrand if the measurable correspondence  $\text{hypo } \Psi : \Omega \rightrightarrows \mathbb{R}^{dT} \times \mathbb{R}$  defined by

$$\text{hypo } \Psi(\omega) = \{(x, y) \in \mathbb{R}^{dT} \times \mathbb{R} \mid \Psi(\omega, x) \geq y\}$$

is closed valued and  $\mathcal{F}$ -measurable in the sense of set-valued maps; see [24, Definitions 14.1 and 14.27]. Note that the correspondence  $\text{hypo } \Psi$  has closed values if and only if the function  $x \mapsto \Psi(\omega, x)$  is upper-semicontinuous for each  $\omega$ ; see [24, Theorem 1.6]. By [24, Corollary 14.34],  $\Psi$  is (jointly) measurable with respect to  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^{dT})$  and  $\mathcal{B}(\overline{\mathbb{R}})$ . Classical examples of normal integrands, which are most prevalent in mathematical finance, are Caratheodory maps; see [24, Example 14.29].

Denote by  $\mathcal{H}$  the set of all  $\mathbb{F}$ -adapted  $\mathbb{R}^d$ -valued processes  $H := (H_0, \dots, H_{T-1})$  with discrete-time index  $t = 0, \dots, T-1$ . Our goal is to study the following control problem:

$$(4) \quad \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P[\Psi(H_0, \dots, H_{T-1})],$$

where  $\Psi : \Omega \times \mathbb{R}^{dT} \rightarrow \mathbb{R}$  is a concave and  $\mathcal{F}$ -measurable normal integrand.

Recall that a function  $f$  from a Borel subset of a Polish space into  $[-\infty, \infty]$  is called lower semianalytic if the set  $\{f < c\}$  is analytic for all  $c \in \mathbb{R}$ ; in particular any Borel function is lower semianalytic. A concave function  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$  is called proper if  $f(x) > -\infty$  for some  $x \in \mathbb{R}^n$ . We refer to [3] and [23, 24] for more details about the theory of lower semianalytic functions and convex analysis, respectively.

The following conditions are in force throughout the paper.

*Assumption 1.* The map  $\Psi : \Omega \times \mathbb{R}^{dT} \rightarrow \mathbb{R} \cup \{-\infty\}$  satisfies the following:

- (1) for every  $\omega \in \Omega$ , the map  $x \mapsto \Psi(\omega, x)$  is concave and upper-semicontinuous;
- (2) there exists a constant  $C \in \mathbb{R}$  such that  $\Psi(\omega, x) \leq C$  for all  $\omega \in \Omega$ ,  $x \in \mathbb{R}^{dT}$ ;
- (3) the map  $(\omega, x) \mapsto \Psi(\omega, x)$  is lower semianalytic;
- (4) there exists  $h^\circ \in \mathbb{R}^{dT}$ , an  $\varepsilon > 0$ , and a constant  $c > 0$  such that

$$(5) \quad \Psi(\omega, x) \geq -c \quad \forall \omega \in \Omega, \forall x \in \mathbb{R}^{dT} : \|x - h^\circ\| \leq \varepsilon.$$

*Remark 2.* At first glance, Assumption 1(2) may look to be rather restrictive. However, it was shown in [18, Example 2.3] that for any (nondecreasing, strictly concave) utility function  $U$  being *unbounded from above*, one can construct a frictionless

market  $S$  and a set  $\mathfrak{P}$  of probability measures such that

$$u(x) := \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P[U(x + H \bullet S_T)] < \infty$$

for any initial capital  $x > 0$ , but there is no maximizer  $\hat{H}$ . So already in the special case of  $\Psi(H) := U(V(H))$  in the frictionless market, the existence may fail for utility functions not being bounded from above.

*Remark 3.* The mapping  $\Psi$  satisfying Assumption 1 is a normal integrand. Indeed, Assumption 1(3) and [3, Lemma 7.29, p. 174] imply that the map  $\omega \mapsto \Psi(\omega, x)$  is  $\mathcal{F}$ -measurable for every  $x \in \mathbb{R}^{dT}$ ; recall that  $\mathcal{F}$  is the universal  $\sigma$ -field on  $\Omega$ . Therefore, normality of  $\Psi$  follows directly from [24, Proposition 14.39, p. 666].

*Remark 4.* Assumption 1(4) is primarily a statement saying that the domain  $\text{dom } \Psi(\omega)$  of the mapping  $x \mapsto \Psi(\omega, x)$  has an interior for each  $\omega \in \Omega$ . It is even stronger, since the strong lower bound on  $\Psi$  on the neighborhood of  $h^\circ$  implies that every strategy  $H \in \mathcal{H}$  satisfying  $\|H(\omega) - h^\circ\| \leq \varepsilon$  for every  $\omega \in \Omega$  satisfies also  $\inf_{P \in \mathfrak{P}} \mathbb{E}^P[\Psi(H)] \geq -c$ . This is a strong regularity condition that plays a crucial role in establishing measurability of the objects arising in the dynamic programming procedure. More precisely, if  $h^\circ = 0$ , it implies that it is enough to know the market on the countable set of deterministic strategies  $H \in \mathbb{Q}^{dT}$ ; cf. Remark 10. It can certainly be relaxed; however, it is difficult to come up with a set of conditions on  $\Psi$  that could be checked a priori. The most general sets of conditions should be stated in terms of objects arising in the dynamic programming procedure; see Remark 23(3) and the proof of Lemma 33.

*Remark 5.* We point out that our definition of a normal integrand  $\Psi$  varies from the classical one in convex optimization as defined, e.g., in [24, Chapter 14], in the sense that  $-\Psi$  is a normal intergrand in classical convex analysis. As we are looking for a maximum of a concave function, our definition of a normal function fits into our setting.

We now define the horizon function  $\Psi^\infty : \Omega \times \mathbb{R}^{dT} \rightarrow \mathbb{R} \cup \{-\infty\}$  of a concave, proper, upper-semicontinuous integrand  $\Psi(\omega, \cdot)$  by

$$\Psi^\infty(\omega, h) = \lim_{n \rightarrow \infty} \frac{1}{n} [\Psi(\omega, x + nh) - \Psi(\omega, x)] = \inf_{n \in \mathbb{N}} \frac{1}{n} [\Psi(\omega, x + nh) - \Psi(\omega, x)],$$

where  $x \in \mathbb{R}^{dT}$  is any vector with  $\Psi(\omega, x) > -\infty$ . Note that for any fixed  $\omega \in \Omega$ , the map  $h \mapsto \Psi^\infty(\omega, h)$  does not depend on the choice of  $x$  in the definition; see [24, Theorem 3.21, p. 87]. The mapping  $\Psi^\infty(\omega, \cdot)$  is positively homogeneous, concave, and upper-semicontinuous; see [24, Theorem 3.21, p. 87]. If, in addition,  $\Psi$  is normal, then so is  $\Psi^\infty$ ; see [24, Exercise 14.54(a), p. 673].

Throughout the paper we impose the following condition:

$$(\text{NA}(\mathfrak{P})) \quad \text{the set } \mathcal{K} := \{H \in \mathcal{H} \mid \Psi^\infty(H_0, \dots, H_{T-1}) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\} \quad \text{is linear.}$$

We call it the no-arbitrage condition. Of course, the set  $\mathcal{H}$  of adapted strategies is a linear space, hence  $\mathcal{K}$  is a subset of it and the definition makes sense. Note that the set  $\mathcal{K}$  is a convex cone; this follows directly from concavity of the map  $\Psi^\infty$ .

*Remark 6.* Naming the condition  $\text{NA}(\mathfrak{P})$  a (robust) no-arbitrage condition is motivated by the following observation: considering a frictionless market with corre-

sponding price process  $S$  and letting

$$\Psi(H) := \sum_{t=0}^{T-1} H_t \cdot (S_{t+1} - S_t) =: H \cdot S_T$$

denote the capital gains from trading in the market using strategy  $H \in \mathcal{H}$ , our notion of  $\text{NA}(\mathfrak{P})$  coincides with the robust no-arbitrage notion in [6, 18]. This follows directly since linearity of  $\Psi$  in this situation implies that  $\Psi^\infty = \Psi$ . Indeed, if  $\mathcal{K}$  were not linear, there would exist a strategy  $H \in \mathcal{K}$  such that  $-H \notin \mathcal{K}$ . But this implies that  $H \cdot S_T \geq 0$   $\mathfrak{P}$ -q.s. and there exists a measure  $P \in \mathfrak{P}$  with  $P[-H \cdot S_T < 0] > 0$ , i.e.,  $P[H \cdot S_T > 0] > 0$ , which means that  $H$  is a robust arbitrage strategy in the sense of [6, 18], and vice versa.

Moreover, when the map  $\Psi$  is of the form initially considered, i.e., given by

$$\Psi(H) = U(V(H)),$$

a sufficient condition (independent of the utility function  $U$ !) for  $\Psi$  to satisfy  $\text{NA}(\mathfrak{P})$  is  $V$  satisfying the following condition:

$$\text{the set } \{H \in \mathcal{H} \mid V^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\} \text{ is linear.}$$

This is a condition on the financial market model, where  $V(H)$  denotes the terminal wealth when investing with strategy  $H$ . Indeed, let  $U$  be a nondecreasing, concave utility function such that  $\Psi$  is not identically equal to  $-\infty$ . Then by Lemma 35

$$\Psi^\infty(h) = \begin{cases} U^\infty(V^\infty(h)) & \text{if } V^\infty(h) > -\infty, \\ -\infty & \text{otherwise.} \end{cases}$$

We claim that linearity of the set  $\{H \in \mathcal{H} \mid V^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\}$  implies linearity of the set  $\{H \in \mathcal{H} \mid \Psi^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\}$ . To see this, observe that  $U^\infty(0) = 0$ , and  $U^\infty$  is nondecreasing as  $U$  is so, too. Let  $H \in \mathcal{H}$  satisfy  $\Psi^\infty(H) \geq 0$   $\mathfrak{P}$ -q.s. By the monotonicity of  $U^\infty$  and as  $\Psi^\infty = U^\infty \circ V^\infty$ , this means that  $V^\infty(H) \geq 0$   $\mathfrak{P}$ -q.s. By assumption on the linearity of the set  $\{H \in \mathcal{H} \mid V^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\}$  we have  $V^\infty(-H) \geq 0$   $\mathfrak{P}$ -q.s., which implies that  $U^\infty(V^\infty(-H)) \geq 0$   $\mathfrak{P}$ -q.s. Linearity of the set  $\{H \in \mathcal{H} \mid \Psi^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\}$  now follows.

Generalized notions of no-arbitrage conditions in the form of linearity-type conditions were already obtained in [20, 21] for markets without uncertainty (i.e., where one fixed measure  $P$  is given). Our no-arbitrage condition  $\text{NA}(\mathfrak{P})$  can be interpreted as an extension of the linearity type of no-arbitrage conditions to the robust framework.

The main theorem of this paper is the following.

**THEOREM 7.** *Let  $\Psi$  be a map satisfying Assumption 1. If the no-arbitrage condition  $\text{NA}(\mathfrak{P})$  holds, then there exists a process  $\hat{H} \in \mathcal{H}$  such that*

$$(6) \quad \inf_{P \in \mathfrak{P}} E^P[\Psi(\hat{H}_0, \dots, \hat{H}_{T-1})] = \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P[\Psi(H_0, \dots, H_{T-1})].$$

We will give the proof of this theorem in section 5.

**3. Examples.** In this section, we give several examples of robust utility maximization in various models of financial markets fitting into the setting of Theorem 7. This was our initial motivation for the abstract robust optimization problem.

*Example 8.* In this example, we analyze the robust utility maximization problem in a classical frictionless market similar to [18].

Let  $S = (S^1, \dots, S^d)$  be a  $d$ -dimensional stock price process with nonnegative components being Borel measurable and constant  $S_0^j = s_0^j > 0$  for all  $j$ . Consider a random utility function  $U: \Omega \times \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ , i.e.,  $U(\omega, \cdot)$  is a nondecreasing, concave function, which is upper-semicontinuous and bounded from above by a constant. Moreover, assume that  $(\omega, y) \mapsto U(\omega, y)$  is lower semianalytic and  $\omega \mapsto U(\omega, y)$  is bounded from below for each  $y > 0$ ; the last two conditions are trivially satisfied if  $U: \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$  is a classical utility function independent of  $\omega$  satisfying  $U(y) > -\infty$  for  $y \in (0, \infty)$ . We define the mapping  $\Psi$  by

$$\Psi(H) = U(x + H \cdot S_T),$$

where  $x > 0$  is the fixed initial wealth of the trader. We want to show that the  $\Psi$  we just defined satisfies Assumption 1. To that end, note that it is concave as  $\Psi(\omega, \cdot)$  is a compositum of a concave and a linear function. Also upper-semicontinuity is clear as  $U$  is upper-semicontinuous and  $H \mapsto H \cdot S_T$  is continuous for every  $\omega$ . As the utility function is bounded from above, the same holds for the mapping  $\Psi$ . Due to the assumption on  $U$  being lower semianalytic, the same holds true for  $\Psi$  being a precomposition of a lower semianalytic function with a Borel function; see [3, Lemma 7.30(3), p. 177].

To see that Assumption 1(4) is satisfied, set

$$\rho := \frac{x}{2dT \max\{s_0^j : j = 1, \dots, d\}}$$

and define the deterministic strategy  $h_t^\circ := (h_t^{\circ,1}, \dots, h_t^{\circ,d}) \in \mathbb{R}^{dT}$  by

$$h_t^{\circ,j} := (T - t)\rho \quad \text{for } t = 0, \dots, T-1, \quad j = 1, \dots, d.$$

Let  $\varepsilon < \frac{\rho}{3}$ . It is easy to see that any  $z \in \mathbb{R}^{dT}$  satisfying  $\|z - h^\circ\| \leq \varepsilon$  is decreasing in each of the components; i.e., the (deterministic) process  $(z_t^j)_{t \in \{0, \dots, T-1\}}$  is decreasing for each  $j = 1, \dots, d$ . We claim that the corresponding capital gains  $z \cdot S_T$  at time  $T$  are uniformly bounded from below. To see this, let  $z \in \mathcal{H}$  be one of such (deterministic) strategies. Writing  $z_{-1} = 0 = z_T$ , we obtain the corresponding capital gains

$$z \cdot S_T = \sum_{t=0}^{T-1} z_t \cdot (S_{t+1} - S_t) = - \sum_{t=0}^T (z_t - z_{t-1}) \cdot S_t.$$

By assumption on nonnegativity of each stock price process  $S^j$  and the fact that the strategy  $(z_t^j)_{t \in \{0, \dots, T-1\}}$  is decreasing, it follows that  $-\sum_{t=1}^T (z_t - z_{t-1}) \cdot S_t \geq 0$ . To see that also  $-(z_0 - z_{-1}) \cdot S_0$  is bounded from below, we use the definition of  $\rho$  to see that

$$-(z_0 - z_{-1}) \cdot S_0 = - \sum_{j=1}^d (z_0^j - z_{-1}^j) S_0^j = - \sum_{j=1}^d z_0^j S_0^j \geq - \sum_{j=1}^d \rho \left( \frac{1}{3} + T \right) s_0^j > -x.$$

Thus, the claim holds true, i.e., for some constant  $\delta$ ,  $x + z \cdot S_T \geq \delta > 0$  for each  $(z_t)$  satisfying  $\|z - h^\circ\| \leq \varepsilon$  and each  $\omega$ . Therefore,  $\Psi$  satisfies Assumption 1(4), as we assumed that  $\omega \mapsto U(\omega, \delta)$  is bounded from below.

*Remark 9.* Our main theorem is more general than the result from [18] in a few directions. First, we do not assume that the stock price process is adapted and there-



fore also include, e.g., the setup of [14]. Moreover, we do not impose the assumption that the utility function is defined on the positive half-line. This allows one to include, for instance, the exponential utility function.

*Remark 10.* One models portfolio constraints by correspondences  $D_t: \Omega^t \rightrightarrows \mathbb{R}^d$  requiring that  $H_t(\omega) \in D_t(\omega)$  for all  $\omega, t$ . Convex constraints  $D_t: \Omega^t \rightrightarrows \mathbb{R}^d$ , which are given by Borel measurable, closed-valued correspondences, are included in our model as long as either  $D_t = \{h^\circ\}$  or  $D_t$  has an interior in such a way that Assumption 1(4) is satisfied. However, one can read out from the proof of our main theorem that Assumption 1(4) can be relaxed for constraints, i.e., it is not necessary for the sets  $D_t$  to include a ball around  $h^\circ$ , and it is enough for them to have an open interior for the dynamic programming with the criterion function

$$\widehat{\Psi}(\omega, h) = \Psi(\omega, h) - \sum_{t=0}^{T-1} \chi_{D_t(\omega)}(h_t)$$

to give the existence of a maximizer; here  $\chi_A$  denotes the convex analytic indicator function for the set  $A \equiv D_t(\omega)$  giving the value 0 if  $h_t \in A$  and  $\infty$  otherwise.

*Example 11.* In this example, we consider the financial market from [10] with proportional transaction costs. The mark-to-market value of the portfolio strategy  $H$  with initial (fixed) capital  $x > 0$  is defined by

$$V(H) = x + \sum_{t=0}^{T-1} H_t \cdot (S_{t+1} - S_t) - \kappa S_t |H_t - H_{t-1}|,$$

where we set  $H_{-1} = 0$ . The one-dimensional stock price process  $(S_t)$  is assumed to be Borel measurable and nonnegative starting at a constant  $S_0 = s_0 > 0$ . The constant  $1 > \kappa \geq 0$  indicates the amount of transaction costs. One then defines  $\Psi$  to be

$$\Psi(\omega, H) = U(V(H))$$

for a random utility function  $U: \Omega \times \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$  being defined as in Example 8. It is easy to check that the conditions of Assumption 1 are satisfied. Concavity and upper-semicontinuity are clear by the fact that  $U$  is upper-semicontinuous, concave, and nondecreasing and  $V$  is concave and continuous. Boundedness from above is clear by the same assumption on the utility function  $U$ . The lower semianalyticity of  $\Psi$  is fulfilled as  $S$  is assumed to be Borel. Now, rewrite the value of the strategy as follows:

$$\begin{aligned} V(H) &= x - \sum_{t=0}^{T-1} S_t (\kappa |H_t - H_{t-1}| + (H_t - H_{t-1})) + H_{T-1} S_T \\ &= x - \sum_{t=0}^{T-1} S_t f(H_t - H_{t-1}) + H_{T-1} S_T, \end{aligned}$$

where the function  $f: h \mapsto \kappa |h| + h$  is less than or equal to zero for all  $h \leq 0$  by assumption on the constant  $\kappa$ . Define  $(h_t^\circ)_{t=0, \dots, T-1} \in \mathbb{R}^T$  and  $\rho$  as in Example 8, and choose  $\varepsilon < \rho \min\{\frac{1}{3}, \frac{T(1-\kappa)}{1+\kappa}\}$ . We know from Example 8 that any  $H \in \mathcal{H}$  satisfying  $\|H - h^\circ\| \leq \varepsilon$  is positive and decreasing, hence

$$- \sum_{t=1}^{T-1} S_t f(H_t - H_{t-1}) + H_{T-1} S_T \geq 0.$$

Moreover, as  $\varepsilon < \rho \frac{T(1-\kappa)}{1+\kappa}$ , it is straightforward to see that

$$-S_0 f(H_0 - H_{-1}) = -s_0(\kappa H_0 + H_0) \geq -s_0(T\rho + \varepsilon)(1 + \kappa) > -x.$$

Therefore, there exists a constant  $\delta$  such that  $V(H) \geq \delta > 0$  for all  $\omega$  and all  $H$  satisfying  $\|H - h^\circ\| \leq \varepsilon$ . Hence  $\Psi = U \circ V$  satisfies Assumption 1(4), as we assumed that  $\omega \mapsto U(\omega, \delta)$  is bounded from below.

*Remark 12.* In Remark 6 we motivated the name “no-arbitrage” for the condition  $\text{NA}(\mathfrak{P})$  by indicating that it coincides with the definition of robust no-arbitrage condition introduced in [6]; we sketch a similar argument for the case of proportional transaction costs. Questions of no-arbitrage in models with proportional transaction costs are usually addressed in the setting of [13]. The fundamental contribution to the no-arbitrage theory in the setup without model uncertainty was made in [25], where the concept of robust no-arbitrage was introduced. More precisely, it was shown that its robust no-arbitrage condition is equivalent to linearity of a certain set of portfolio rebalancings; see [25, Lemma 2.6]. One can translate the example above to the framework of [13]. Observe that our condition  $\text{NA}(\mathfrak{P})$  in fact generalizes the robust no-arbitrage condition of [25], i.e., an equivalent formulation thereof, to this nondominated setup.

*Remark 13.* The example above treated the mark-to-market value in the market. One could equivalently consider the liquidation value of the portfolio  $H$ . Also, it is easy to see that this is not restricted to proportional transaction costs but can be extended to more general transaction costs. The important thing is the condition Assumption 1(4) and the example above indicates where it comes up in the argument.

*Remark 14.* One could also consider (proportional) transaction costs in physical units. We will only sketch this example. There are  $d$  risky assets in the market and the portfolio of the trader is described by specifying at each time  $t$  the number of shares in each of the  $d$  risky assets. Going back to the original contribution of [13], one models trading in the market by specifying in each time instance  $t = 0, \dots, T-1$  how many shares  $H_t^{\alpha \rightarrow \beta}$  of risky asset  $\alpha$  to transfer to shares of the risky asset  $\beta$ . Hence, a strategy will be a matrix with adapted entries. The market mechanism, i.e., changes of the portfolio due to a trade order, is given by a sequence of maps

$$F_t : \Omega^t \times \mathbb{R}^{d \times d} \longrightarrow \mathbb{R}^d.$$

The interpretation is the following. After executing the order  $H$  at time  $t$ , the holdings of the trader in physical units are going to change by  $F_t(H)$ . So, the portfolio of the trader at the end of the trading period is given by

$$V(H) = \sum_{t=0}^{T-1} F_t(H_t).$$

We also refer to [5].

Now, let the utility function be denoted  $U : \mathbb{R}^d \rightarrow \mathbb{R}$ , and define the map  $\Psi$  by

$$\Psi(H) = U(V(H)).$$

One can easily find conditions under which Assumption 1 is satisfied: for instance, the utility function  $U$  is bounded from above, concave, nondecreasing, and upper-semicontinuous with respect to partial order given by an order cone  $\mathbb{R}_+^d$ ; the market

impact functions  $F_t$  are continuous and concave with respect to the same partial order; appropriate measurability conditions and boundedness conditions are used to obtain conditions Assumption 1(3), (4).

*Example 15.* In this example, we consider the optimal liquidation problem, an adaptation of the model introduced in [1]. The agent initially holds  $X > 0$  units of a security, which need to be liquidated by the time of maturity  $T$ . The strategy  $H = (H_0, \dots, H_{T-1})$  denotes that the agent holds  $H_t$  units of the security after liquidation of the amount  $H_{t-1} - H_t$  at time  $t$ ; in the interpretation of [1] the agent liquidates this amount between times  $t-1$  and  $t$ . This liquidation yields  $(H_{t-1} - H_t)\hat{S}_t$  on the bank account, where  $\hat{S}_t$  encodes the price achieved per unit of security, given the temporary price impact of trading. The temporary price impact is modeled by the function  $\mathbf{g}$ , writing  $\hat{S}_t = S_t - \mathbf{g}(H_{t-1} - H_t)$ , where  $S_t$  would denote the stock price at time  $t$  where there is no trading in the market. We set  $H_{-1} = X$  and  $H_T = 0$ . After liquidation the amount on the bank account  $V(H)$  is given by

$$\begin{aligned} V(H) &:= \sum_{t=0}^T (H_{t-1} - H_t) \hat{S}_t \\ &= \sum_{t=0}^T (H_{t-1} - H_t) S_t - \sum_{t=0}^T (H_{t-1} - H_t) \mathbf{g}(H_{t-1} - H_t) \\ &= X S_0 + \sum_{t=0}^{T-1} H_t (S_{t+1} - S_t) - \sum_{t=0}^T (H_{t-1} - H_t) \mathbf{g}(H_{t-1} - H_t). \end{aligned}$$

We assume that the function  $\mathbf{g}$  is convex, lower-semicontinuous, and nondecreasing and satisfies  $\mathbf{g}(0) = 0$ . Moreover, we assume that  $\mathbf{g}(x) < \infty$  for all  $x \geq 0$ . Then the above defines a concave market model. The first term above does not depend on the strategy, the second one depends linearly, and the last term is obviously concave; indeed, the function  $x \mapsto x\mathbf{g}(x)$  is, under stated conditions, clearly convex on  $x \geq 0$ .

Let  $U: \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$  be a nondecreasing and upper-semicontinuous, concave utility function bounded from above. Define the map  $\Psi$  by  $\Psi(H) := U(V(H))$ . Clearly,  $\Psi(\omega, \cdot)$  is concave and upper-semicontinuous as  $V(\cdot)$  is concave, upper-semicontinuous, and  $U$  is concave, upper-semicontinuous and nondecreasing. Boundedness from above follows from the same assumption on  $U$ , and Borel measurability holds whenever the stock price process  $(S_t)$  is Borel.

We assume that the stock price process  $(S_t)$  is bounded below by some constant  $\gamma \leq 0$ , i.e.,  $S_t(\omega) \geq \gamma$  for all  $\omega$ . We then show that Assumption 1(4) is satisfied by the model. Define the sequence that encodes constant rate of liquidation

$$h_t^\circ := X \left( 1 - \frac{t+1}{T+1} \right), \quad t \in \{0, \dots, T-1\},$$

and choose  $\varepsilon < \frac{X}{3(T+1)}$ . Then, it is straightforward to see that for any  $(x_t)$  satisfying  $\|x - h^\circ\| < \varepsilon$ , setting also  $x_{-1} = X$  and  $x_T = 0$ , we have

$$\begin{aligned} V(x) &= \sum_{t=0}^T (x_{t-1} - x_t) S_t - \sum_{t=0}^{T-1} (x_{t-1} - x_t) \mathbf{g}(x_{t-1} - x_t) \\ &\geq \frac{5\gamma X}{3(T+1)} - \frac{5X}{3(T+1)} \mathbf{g}\left(\frac{5X}{3(T+1)}\right). \end{aligned}$$

Therefore,  $\Psi$  satisfies Assumption 1.

**4. One-period-model.** The key technique of the proof of Theorem 7 is dynamic programming, i.e., backward induction, where one optimizes the strategy over each single time step and then “glues” the single step strategies together. To that end, it is necessary to analyze the corresponding one-period model first, which we do in this section. In the one-step case one may prove the result in a more general setup than the multiperiod case. For that reason we first provide the setup for the one-period case.

**4.1. Setup.** Let  $\Omega$  be a Polish space,  $\mathcal{F}$  be the universal completion of the Borel  $\sigma$ -field  $\mathcal{B}(\Omega)$ , and  $\mathfrak{P}$  be a possibly nondominated set of probability measures on  $\mathcal{F}$ .

We fix a function  $\Psi: \Omega \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$  and consider the optimization problem

$$(7) \quad \sup_{h \in \mathbb{R}^d} \inf_{P \in \mathfrak{P}} E^P[\Psi(h)].$$

Throughout this section, we will work with the following conditions on  $\Psi$ .

*Assumption 16.* The map  $\Psi: \Omega \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$  satisfies the following:

- (1) for every  $\omega \in \Omega$ , the map  $x \mapsto \Psi(\omega, x)$  is concave and upper-semicontinuous;
- (2) there exists a constant  $C \in \mathbb{R}$  such that  $\Psi(\omega, x) \leq C$  for all  $\omega \in \Omega, x \in \mathbb{R}^d$ ;
- (3) the map  $(\omega, x) \mapsto \Psi(\omega, x)$  is lower semianalytic;
- (4) the optimization problem is nontrivial, i.e., there exists a strategy  $h^\circ \in \mathcal{H} = \mathbb{R}^d$ , such that  $\inf_{P \in \mathfrak{P}} E_P[\Psi(h^\circ)] > -\infty$ .

*Remark 17.* Let us compare Assumption 16 with Assumption 1 of the general multiperiod case. First note that points (1), (2), and (3) are the same in both cases. In fact, Assumption 16(3) ensures that for every  $x \in \mathbb{R}^d$  the random variable  $\Psi(x)$  is measurable. This can be achieved, e.g., by assuming the weaker condition of  $(\omega, x) \mapsto \Psi(\omega, x)$  is universally measurable; see [3, Lemma 7.29, p. 174]. However in the multiperiod case, lower semianalyticity cannot be relaxed without losing the measurability setting needed to be able to apply crucial measurable selection arguments; see the proof of Propositions 26 and 28. Assumption 16(4) just requires that the optimization problem is well posed, i.e., the value is not equal to  $-\infty$  for all  $x \in \mathbb{R}^d$ . The stronger Assumption 1(4) serves a different purpose; see Remark 23(3).

We work under the following no-arbitrage condition:

$$(8) \quad \text{the set } \mathcal{K} := \{h \in \mathbb{R}^d \mid \Psi^\infty(h) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\} \text{ is linear.}$$

The main result of this section is the existence of a maximizer for the one-period optimization problem (7).

**THEOREM 18.** *Let the no-arbitrage condition (8) and Assumption 16 hold. There exists a strategy  $\hat{h} \in \mathbb{R}^d$  such that*

$$(9) \quad \inf_{P \in \mathfrak{P}} E_P[\Psi(\hat{h})] = \sup_{h \in \mathbb{R}^d} \inf_{P \in \mathfrak{P}} E_P[\Psi(h)],$$

i.e., the optimization problem admits an optimizer.

**4.2. Proof of the one-period optimization problem.** In the one-step case, one can prove the existence of a maximizer just by applying classical arguments from convex analysis. Assumption 16 is in force throughout this section. One of the key elements of the proof is the following reformulation of [23, Theorem 9.2, p. 75].

PROPOSITION 19. Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$  be concave, upper-semicontinuous, and proper, and let  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be linear. If the set  $\{z \in \mathbb{R}^n \mid Az = 0, f^\infty(z) \geq 0\}$  is linear, then the function

$$g(y) := \sup\{f(x) \mid x \in \mathbb{R}^n, Ax = y\}$$

is concave, proper, and upper-semicontinuous. Moreover,

$$g^\infty(y) = \sup\{f^\infty(x) \mid x \in \mathbb{R}^n, Ax = y\}.$$

Furthermore, for each  $y$  such that  $g(y) > -\infty$ , the supremum in the definition of  $g$  is attained.

Consider the function

$$\Phi : h \mapsto \inf_{P \in \mathfrak{P}} E^P[\Psi(h)].$$

Observe that  $\Phi$  is concave and upper-semicontinuous as an infimum of concave upper-semicontinuous functions; use Assumption 16(1), (2) and Fatou's lemma. It is also proper, i.e., not identically equal to  $-\infty$ , by Assumption 16(4). Moreover, we have the following.

LEMMA 20. Let  $\Psi : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$  satisfy Assumption 16. Then

$$\Phi^\infty(h) = \inf_{P \in \mathfrak{P}} E^P[\Psi^\infty(h)]$$

for all  $h \in \mathbb{R}^d$ .

*Proof.* By concavity of  $\Psi(\omega, \cdot)$ , the sequence  $n \mapsto \frac{1}{n}(\Psi(\omega, nh + h^\circ) - \Psi(\omega, h^\circ))$  is pointwise decreasing for every  $\omega$ , hence monotone convergence yields

$$\begin{aligned} \Phi^\infty(h) &= \inf_{n \in \mathbb{N}} \frac{1}{n} \left( \inf_{P \in \mathfrak{P}} E^P[\Psi(nh + h^\circ)] - \inf_{P \in \mathfrak{P}} E^P[\Psi(h^\circ)] \right) \\ &= \inf_{n \in \mathbb{N}} \frac{1}{n} \left( \inf_{P \in \mathfrak{P}} (E^P[\Psi(nh + h^\circ) - \Psi(h^\circ)] + E^P[\Psi(h^\circ)]) - \inf_{P \in \mathfrak{P}} E^P[\Psi(h^\circ)] \right) \\ &\geq \inf_{n \in \mathbb{N}} \frac{1}{n} \inf_{P \in \mathfrak{P}} (E^P[\Psi(nh + h^\circ) - \Psi(h^\circ)]) \\ &= \inf_{P \in \mathfrak{P}} \lim_{n \rightarrow \infty} \frac{1}{n} (E^P[\Psi(nh + h^\circ) - \Psi(h^\circ)]) \\ &= \inf_{P \in \mathfrak{P}} E^P[\Psi^\infty(h)]. \end{aligned}$$

For the reverse inequality, we use Assumption 16(2) and come back to the second line of the above calculation to get

$$\begin{aligned} \Phi^\infty(h) &= \inf_{n \in \mathbb{N}} \frac{1}{n} \left( \inf_{P \in \mathfrak{P}} (E^P[\Psi(nh + h^\circ) - \Psi(h^\circ)] + E^P[\Psi(h^\circ)]) - \inf_{P \in \mathfrak{P}} E^P[\Psi(h^\circ)] \right) \\ &\leq \inf_{n \in \mathbb{N}} \frac{1}{n} \left( \inf_{P \in \mathfrak{P}} (E^P[\Psi(nh + h^\circ) - \Psi(h^\circ)]) + C - \inf_{P \in \mathfrak{P}} E^P[\Psi(h^\circ)] \right) \\ &= \inf_{P \in \mathfrak{P}} E^P[\Psi^\infty(h)]. \end{aligned}$$

□

*Proof of Theorem 18.* The mapping  $\Psi(\omega, \cdot)$  is bounded from above by a constant  $C$  for each  $\omega \in \Omega$ . Thus  $\Psi^\infty(\omega, \cdot) \leq 0$  for each  $\omega$ . From Lemma 20 we know that  $\Phi^\infty(h) = \inf_{P \in \mathfrak{P}} E^P[\Psi^\infty(h)]$  and thus also  $\Phi^\infty(h) \leq 0$  for all  $h \in \mathbb{R}^d$ .

Observe that  $\Phi^\infty(h) = 0$  if and only if  $h \in \mathcal{K}$ . Therefore, by the no-arbitrage condition (8), we see that the conditions of Proposition 19 are fulfilled for the linear map  $A : \mathbb{R}^d \rightarrow \{0\}$ ,  $x \mapsto 0$ , and  $f \equiv \Phi$ . Thus, Proposition 19 yields the existence of a maximizer  $\hat{h}$ , as (by choosing  $y \equiv 0$ )

$$\sup_{h \in \mathbb{R}^d} \Phi(h) = \sup\{\Phi(h) \mid Ah = 0\} = \sup_{h \in \mathbb{R}^d} \inf_{P \in \mathfrak{P}} E^P[\Psi(h)] \geq \inf_{P \in \mathfrak{P}} E^P[\Psi(h^\circ)] > -\infty. \quad \square$$

*Remark 21.* Notice that in the above proof we did not require any structural properties of the measurable space  $(\Omega, \mathcal{F})$  or of the set  $\mathfrak{P}$  of probability measures. The only important element of the setup is concavity of the map  $\Psi$ . In particular, Theorem 18 remains valid for any measurable space  $(\Omega, \mathcal{F})$  when replacing Assumption 16(3) by the assumption that the map  $(\omega, x) \mapsto \Psi(\omega, x)$  is  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable.

**5. Multiperiod model.** The key idea of the multiperiod case is to adapt the techniques of the dynamic programming principle, developed in [11], to the robust framework. We will use the setup and measurability techniques developed in [6] and [18].

From this point on, Assumption 1 is in force as well as the no-arbitrage condition  $\text{NA}(\mathfrak{P})$  given by

$$\text{the set } \mathcal{K} := \{H \in \mathcal{H} \mid \Psi^\infty(H) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\} \text{ is linear.}$$

The basic idea of dynamic programming is to reduce the maximization over the set of strategies  $\mathcal{H}$  to one-step maximization we encountered in the previous section. Precisely, using the notation where for  $\omega^t \in \Omega^t$  and  $\tilde{\omega} \in \Omega_1$ ,  $\omega^t \otimes_t \tilde{\omega}$  stands for the pair  $(\omega^t, \tilde{\omega}) \in \Omega^{t+1}$ , we define the following sequences of maps: set  $\Psi_T := \Psi$  and for  $t = T - 1, \dots, 0$  and  $\omega^t \in \Omega^t$  define

$$\begin{aligned} \Phi_t(\omega^t, x^{t+1}) &:= \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P[\Psi_{t+1}(\omega^t \otimes_t \cdot, x^{t+1})], \\ \tilde{\Psi}_t(\omega^t, x^t) &:= \sup_{\tilde{x} \in \mathbb{R}^d} \Phi_t(\omega^t, x^t, \tilde{x}), \\ \Psi_t(\omega^t, \cdot) &:= \text{cl } \tilde{\Psi}_t(\omega^t, \cdot), \end{aligned}$$

where by  $\text{cl } \tilde{\Psi}_t(\omega^t, \cdot)$  we denoted the upper-semicontinuous hull of the function  $x \mapsto \tilde{\Psi}_t(\omega^t, x)$ . The upper-semicontinuous hull of a function  $f : \mathbb{R}^n \rightarrow [-\infty, \infty]$  is the smallest upper-semicontinuous function (not necessarily finite) minorized by  $f$ ; it is the function whose hypograph is the closure in  $\mathbb{R}^n \times \mathbb{R}$  of the hypograph of  $f$ ; see [23, p. 52].

*Remark 22.* Let us give a rough sketch of an argument why one would consider the recursion above. We will restrict our attention to the case with  $T = 2$ . The optimization problem we are considering is

$$\sup_{H_0, H_1} \inf_{P \in \mathfrak{P}} E^P[\Psi_2(H_0, H_1)] = \sup_{H_0} \left[ \sup_{H_1} \inf_{P \in \mathfrak{P}} E^P[\Psi_2(H_0, H_1)] \right].$$

After we evaluated the expression in the brackets, optimization over  $H_0$  will follow the argument we provided for the one-step case. Let us, therefore, concentrate on

the optimization over  $H_1$ . First, calculate the conditional expectation of  $\Psi_2(H_0, H_1)$  given  $\mathcal{F}_1$ . By the definition of the probability measure as  $P = P_0 \otimes P_1$ , we know that

$$\begin{aligned} E^{P_0 \otimes P_1}[\Psi_2(H_0, H_1)|\mathcal{F}_1](\omega_1) &= \int_{\Omega_2} \Psi_2(\omega_1 \otimes \omega_2, H_0, H_1) P_1(\omega_1, d\omega_2) \\ &= E^{P_1(\omega_1)}[\Psi_2(\omega_1 \otimes_1 \cdot, H_0, H_1(\omega_1))], \end{aligned}$$

where the second equality is just change in notation. Hence,  $\Phi_1$  is just

$$\Phi_1(\omega_1, x_0, x_1) = \Phi_1(\omega_1, x^2) = \inf_{P_1 \in \mathfrak{P}_1(\omega_1)} E^{P_1}[\Psi_2(x^2)|\mathcal{F}_1](\omega_1),$$

where the versions of the conditional expectation are defined via kernels  $P_1 \in \mathfrak{P}_1(\omega_1)$ . Coming back to the minimization over  $H_1$ , first use the tower property: for any strategy  $H \in \mathcal{H}$  we have

$$\inf_{P \in \mathfrak{P}} E^P[\Psi_2(H_0, H_1)] = \inf_{P_0 \in \mathfrak{P}_0} E^{P_0} \left[ \inf_{P_1 \in \mathfrak{P}_1(\omega_1)} E^{P_1}[\Psi_2(\omega_1 \otimes_1 \cdot, H_0, H_1)] \right]$$

by the decomposability property. However, one would need to show that the expression in the expectation is measurable. In particular, the above equation implies that

$$\sup_{H_1} \inf_{P \in \mathfrak{P}} E^P[\Psi_2(H_0, H_1)] = \sup_{H_1} \inf_{P \in \mathfrak{P}} E^P[\Phi_1(\cdot, H_0, H_1)].$$

It remains to show that

$$\sup_{H_1} \inf_{P \in \mathfrak{P}} E^P[\Phi_1(\cdot, H_0, H_1)] = \inf_{P_0 \in \mathfrak{P}_0} E^{P_0}[\Psi_1(\cdot, H_0)].$$

The inequality  $\leq$  is easy to see. Indeed, we have  $\Phi_1(\omega, H_0, H_1) \leq \Psi_1(\omega, H_0)$  for each  $\omega, H_0$ . To see the converse, it is enough to prove that the supremum in the definition of  $\Psi_1$  is attained by some  $\mathcal{F}_1$ -measurable random variable.

This formally derives the recursion we are considering for the maximization problem.

Of course, the remark above does not constitute a proof to our main theorem but is merely a sketch of an argument for why the recursion we are observing makes sense. Indeed, the main part of the argument will be proving measurability and attainment of maximizers in the one-step case.

*Remark 23.* Before going to the more difficult part of the proof, let us make a few simple observations about the recursion defined above. The numbering refers to the numbering in Assumption 1.

- (1a) It follows directly from the definitions that the mappings  $\Phi_t$  and  $\Psi_t$  are concave. Indeed, we assume that  $\Psi_T = \Psi$  is concave, and [24, Proposition 2.9(b), p. 43] implies that  $\Phi_t$  is concave as soon as  $\Psi_{t+1}$  is. Moreover, [24, Proposition 2.22(a), p. 51] implies that  $\tilde{\Psi}_t$  and hence also  $\Psi_t$  is concave as soon as  $\Phi_t$  is. This holds for each  $\omega^t \in \Omega^t$ .
- (1b) For each  $\omega^t$ , the map  $\Phi_t$  is upper-semicontinuous in the second variable as long as  $\Psi_{t+1}$  is. The argument, which uses the Fatou lemma and the uniform upper bound, is given in the one-step case. Moreover, upper-semicontinuity of  $\Psi_t$  follows directly from its definition as an upper-semicontinuous hull of

the proper concave function  $\tilde{\Psi}_t$ . In Proposition 27(i), we will show that also the function  $\tilde{\Psi}_t$  is upper-semicontinuous for  $\mathfrak{P}$ -q.e.  $\omega^t$ . This will then imply that  $\tilde{\Psi}_t$  coincides  $\mathfrak{P}$ -q.s. with  $\Psi_t$ . The proof will be based on Proposition 19 and a local no-arbitrage condition which we will introduce in Definition 25.

- (2) Boundedness from above by a constant  $C \in \mathbb{R}$  is obvious: take the same constant  $C$  as in Assumption 1(2) for the mapping  $\Psi$ .
- (3) We will prove in Lemma 30 that  $\Phi_t$  is lower semianalytic if  $\Psi_{t+1}$  is. Semianalyticity of  $\Psi_t$  is based on a density argument; see Lemma 32. The problem arises as the supremum of lower semianalytic functions need not be lower semianalytic. Passing to a supremum of a countable number of functions, i.e., maximizing over  $\mathbb{Q}^d$  instead of  $\mathbb{R}^d$  in the definition of  $\tilde{\Psi}_t$ , preserves this measurability property. This is the reason for condition (4) in Assumption 1.
- (4) It is easy to see that condition (5) is satisfied for the mapping  $\Phi_t$  as long as it is satisfied for the mapping  $\Psi_{t+1}$ . To see that the mapping  $\tilde{\Psi}_t$  needs to satisfy (5) when the condition is satisfied by  $\Phi_t$ , it is enough to notice that  $\tilde{\Psi}_t(\omega^t, x^t) \geq \Phi_t(\omega^t, x^t, h_t^\circ)$  for each  $(\omega^t, x^t) \in \Omega^t \times \mathbb{R}^{dt}$ , where  $h_t^\circ$  is the  $t+1$ -entry of the vector  $h^\circ = (h_0^\circ, \dots, h_{T-1}^\circ)$  defined in Assumption 1(4). Then, as  $\Psi_t \geq \tilde{\Psi}_t$  by definition of the upper-semicontinuous hull,  $\Psi_t$  satisfies (5).

*Remark 24.* We just argued in Remark 23 that the mapping  $x^t \mapsto \tilde{\Psi}_t(\omega^t, x^t)$  is concave for each  $t$  and all  $\omega^t \in \Omega^t$  and that  $\Psi_t(\omega^t, h^{\circ,t}) > -c$  for all  $\omega^t$ , where  $h^{\circ,t} \in \mathbb{R}^{dt}$  denotes the restriction of  $h^\circ$  defined in Assumption 1(4) to the first  $t$  entries. Therefore, [24, Theorem 2.35, p. 59] provides the identity

$$\Psi_t(\omega^t, x^t) = \lim_{\lambda \nearrow 1} \tilde{\Psi}_t(\omega^t, \lambda x^t + (1-\lambda)h^{\circ,t}).$$

We first need to show that if the mapping  $\Psi$  satisfies Assumption 1, then also all of the mappings  $\Psi_t$  do for each  $t = 0, \dots, T-1$ . This has to be deduced from the dynamic programming recursion. For that we first need some more terminology and notation. For each  $t \in \{1, \dots, T\}$ , denote by  $\mathcal{H}^t$  the set of all  $\mathbb{F}$ -adapted,  $\mathbb{R}^d$ -valued processes  $H^t := (H_0, \dots, H_{t-1})$ ; these are just restrictions of strategies in  $\mathcal{H}$  to the first  $t$  time steps. Define the no-arbitrage condition up to time  $t$ , denoted by  $\text{NA}(\mathfrak{P})^t$ , for the mappings  $(\Psi_t)$  in the natural way, by saying

the set  $\mathcal{K}^t := \{H^t \in \mathcal{H}^t \mid \Psi_t^\infty(H^t) \geq 0 \text{ } \mathfrak{P}\text{-q.s.}\}$  is linear.

Condition  $\text{NA}(\mathfrak{P})^t$  is a statement about a set of strategies and as such cannot yet be used to prove things that we need it for. What we need is a local version of the no-arbitrage condition.

**DEFINITION 25.** For each  $t \in \{0, \dots, T-1\}$  and  $\omega^t \in \Omega^t$  define a set

$$(10) \quad K_t(\omega^t) := \{h \in \mathbb{R}^d \mid \Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, 0, \dots, 0, h) \geq 0 \text{ } \mathfrak{P}_t(\omega^t)\text{-q.s.}\}.$$

We say that condition  $\text{NA}_t$  holds if

the set  $\{\omega^t \in \Omega^t \mid K_t(\omega^t) \text{ is linear}\}$  has  $\mathfrak{P}$ -full measure.

To make a simple observation, let  $\hat{h}$  be a universally measurable selection of  $K_t$ . Then the strategy  $(0, \dots, 0, \hat{h}) \in \mathcal{K}^{t+1}$ , where there are  $t$  zeros in the previous expression. Thus, it is clear that linearity of  $K_t$  is necessary for the no-arbitrage condition  $\text{NA}(\mathfrak{P})^{t+1}$ .



PROPOSITION 26. Let  $t \in \{0, \dots, T-1\}$ . Assume that  $\Psi_{t+1}$  satisfies Assumption 1. If  $\mathcal{K}^{t+1}$  is linear, then the local no-arbitrage condition  $\text{NA}_t$  holds.

Having this result at hand, one can proceed as in the one-step case. The following is a direct consequence.

PROPOSITION 27. Let  $t \in \{0, \dots, T-1\}$ . Assume that  $\Psi_{t+1}$  satisfies Assumption 1 and that  $\text{NA}_t$  holds. Then

- (i) the map  $x^t \mapsto \tilde{\Psi}_t(\omega^t, x^t)$  is upper-semicontinuous  $\mathfrak{P}$ -q.s.; in particular, it coincides with  $x^t \mapsto \Psi_t(\omega^t, x^t)$  for  $\mathfrak{P}$ -quasi every  $\omega^t$ ;
- (ii) for every  $H^t \in \mathcal{H}^t$  there exists an  $\mathcal{F}_t$ -measurable mapping  $\hat{h}_t : \Omega^t \rightarrow \mathbb{R}^d$  such that

$$\Phi_t(\omega^t, H^t(\omega^t), \hat{h}_t) = \Psi_t(\omega^t, H^t(\omega^t)) \quad \text{for } \mathfrak{P}\text{-q.e. } \omega^t \in \Omega^t.$$

The important step toward the proof of our main result is the observation that sets  $\mathcal{K}^t$  behave well under the dynamic programming recursion.

PROPOSITION 28. Let  $t \in \{0, \dots, T-1\}$  and let  $\Psi_{t+1}$  satisfy Assumption 1. If  $\mathcal{K}^{t+1}$  is linear, then so is  $\mathcal{K}^t$ .

The proof of the following proposition will be done by backward induction.

PROPOSITION 29. For any  $t \in \{0, \dots, T-1\}$ , the function  $\Psi_{t+1}$  satisfies Assumption 1 and the local no-arbitrage condition  $\text{NA}_t$  holds.

The proofs of Propositions 26–29 will be given in the next subsection.

**5.1. Proofs of Propositions 26–29.** This is the technical part of this paper and therefore is divided into several lemmas. We first start with a useful lemma providing the relation between  $\Psi_{t+1}$  and  $\Phi_t$ .

LEMMA 30. Let  $t \in \{0, \dots, T-1\}$ . If  $\Psi_{t+1}$  satisfies Assumption 1, then so does  $\Phi_t$  and for all  $(\omega^t, x^{t+1}) \in \Omega^t \times \mathbb{R}^{d(t+1)}$  we have

$$(11) \quad \Phi_t^\infty(\omega^t, x^{t+1}) = \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P[\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, x^{t+1})].$$

*Proof.* Conditions (1), (2), and (4) are clear by definition and were argued in Remark 23. To see that  $(\omega^t, x^{t+1}) \mapsto \Phi_t(\omega^t, x^{t+1})$  is lower semianalytic, we first recall that the map  $(\omega^{t+1}, x^{t+1}) \mapsto \Psi_{t+1}(\omega^{t+1}, x^{t+1})$  is lower semianalytic by assumption. Also, the map

$$\mathfrak{M}_1(\Omega) \times \Omega^t \times \Omega_1 \times \mathbb{R}^{d(t+1)} \rightarrow \overline{\mathbb{R}}, \quad (P, \omega^t, \tilde{\omega}, x^{t+1}) \mapsto \Psi_{t+1}(\omega^t \otimes_t \tilde{\omega}, x^{t+1})$$

is lower semianalytic as it is independent of the variable  $P$ . Consider the Borel measurable stochastic kernel  $\kappa$  on  $\Omega_1$  given  $\mathfrak{M}_1(\Omega) \times \Omega^t \times \mathbb{R}^{d(t+1)}$  defined by

$$((P, \omega^t, x^{t+1}), A) \mapsto \kappa(A | P, \omega^t, x^{t+1}) := P[A];$$

Borel measurability of the kernel follows from [3, Proposition 7.26, p. 134] and [3, Corollary 7.29.1, p. 144]. Then, applying [3, Proposition 7.48, p. 180] to  $\kappa$ , we obtain that

$$(12) \quad \mathfrak{M}_1(\Omega) \times \Omega^t \times \mathbb{R}^{d(t+1)} \rightarrow \overline{\mathbb{R}}, \quad (P, \omega^t, x^{t+1}) \mapsto E^P[\Psi_{t+1}(\omega^t \otimes_t \cdot, x^{t+1})]$$

is lower semianalytic. By assumption, the graph of  $\mathfrak{P}_t$  is analytic. Therefore, we deduce from [3, Lemma 7.47, p. 179] that

$$(\omega^t, x^{t+1}) \mapsto \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P [\Psi_{t+1}(\omega^t \otimes_t \cdot, x^{t+1})] = \Phi_t(\omega^t, x^{t+1})$$

is lower semianalytic. Finally, by the same arguments as in the proof of Lemma 20, we see directly that (11) holds true.  $\square$

Before we can start with the proof of Proposition 26, we need to see that the set-valued map  $K_t(\omega^t)$  has some desirable properties.

**LEMMA 31.** *Let  $t \in \{0, \dots, T-1\}$ . Assume that  $\Psi_{t+1}$  satisfy Assumption 1. Then the set-valued map  $K_t$  defined in (10) is a closed, convex,  $\mathcal{F}_t$ -measurable correspondence and the set  $\{\omega^t \in \Omega^t \mid K_t(\omega^t) \text{ is linear}\} \in \mathcal{F}_t$ .*

*Proof.* As  $\Psi_{t+1}^\infty$  is concave, positively homogeneous, and upper-semicontinuous in  $x^t$ , and  $K_t(\omega^t)$  is a closed valued convex cone for every  $\omega^t$ . Observe that

$$K_t(\omega^t) = \{x_t \in \mathbb{R}^d \mid \Phi_t^\infty(\omega^t, 0, \dots, 0, x_t) \geq 0\}.$$

By Lemma 30 and Remark 3,  $\Phi_t$  is a concave  $\mathcal{F}_t$ -normal integrand, hence so is  $\Phi_t^\infty$ . Thus, the set-valued map  $K_t$  is an  $\mathcal{F}_t$ -measurable correspondence; see [24, Proposition 14.33, p. 663] and [24, Proposition 14.45(a), p. 669].

Finally, from  $K_t(\omega^t)$  being a convex cone, we get that

$$\{\omega^t \in \Omega^t \mid K_t(\omega^t) \text{ is linear}\} = \{\omega^t \in \Omega^t \mid K_t(\omega^t) = -K_t(\omega^t)\}.$$

By [24, Theorem 14.5(a), p. 646],  $K_t$  admits a Castaing representation  $\{x_n\}$ . Then, we see that

$$\{\omega^t \in \Omega^t \mid K_t(\omega^t) \text{ is linear}\} = \bigcap_{n \in \mathbb{N}} \{\omega^t \mid -x_n(\omega^t) \in K_t(\omega^t)\}.$$

That the latter set is  $\mathcal{F}_t$ -measurable now follows from [24, Definition 14.3(c), p. 644] and [24, Proposition 14.11(c), p. 651].  $\square$

Now we prove that if  $\Psi_{t+1}$  satisfies Assumption 1, linearity of  $\mathcal{K}^{t+1}$  implies  $\text{NA}_t$ .

*Proof of Proposition 26.* Assume that  $\text{NA}_t$  does not hold. Then by definition, there exists a probability measure in  $\mathfrak{P}$  with its restriction to  $\Omega^t$  denoted by  $P^t$  such that the complement of the set

$$G_t := \{\omega^t \in \Omega^t \mid K_t(\omega^t) = -K_t(\omega^t)\}$$

satisfies  $P^t[G_t^c] > 0$ . We claim that there is a strategy of the form  $H^{t+1} = (0, \dots, 0, h_t)$  and a measure  $\tilde{P} \in \mathfrak{P}$  such that  $H^{t+1} \in \mathcal{K}^{t+1}$ , but  $\tilde{P}[\Psi_{t+1}^\infty(-H^{t+1}) < 0] > 0$ , i.e.,  $-H^{t+1} \notin \mathcal{K}^{t+1}$ . The measure  $\tilde{P}$ , restricted to  $\Omega^{t+1}$ , will be defined as  $P^{t+1} = P^t \otimes P_t$  for some selection  $P_t$  of  $\mathfrak{P}_t$ .

*Step 1.* We prove that there is a Borel measurable set-valued map  $K_t^{P^t} : \Omega^t \rightrightarrows \mathbb{R}^d$  that coincides with  $K_t$  for  $P^t$ -a.a.  $\omega^t$ .

Let  $\{x_n\}$  be the Castaing representation of  $K_t$ . By [3, Lemma 7.27, p. 173] we can modify each of these universally measurable selections  $x_n$  on a  $P^t$ -nullset to get almost sure selections  $x_n^{P^t}$  that are Borel measurable. Define a new set-valued map

$$K_t^{P^t}(\omega^t) = \overline{\{x_n^{P^t}(\omega^t) \mid n \in \mathbb{N}\}}.$$

The set-valued map  $K_t^{P^t}$  is Borel measurable by definition and  $K_t^{P^t}(\omega^t) = K_t(\omega^t)$  for  $P^t$ -a.a.  $\omega^t$ ; this follows from [24, Proposition 14.11, p. 651] and [24, Proposition 14.2, p. 644].

*Step 2.* Define the set  $\mathcal{S}_t \subset \Omega^t \times \mathbb{R}^d \times \mathfrak{M}_1(\Omega_1)$  by

$$\mathcal{S}_t := \left\{ (\omega^t, h, P) \mid h \in K_t^{P^t}(\omega^t), P \in \mathfrak{P}_t(\omega^t), E^P [\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, 0, \dots, 0, -h)] < 0 \right\}.$$

We claim that this set  $\mathcal{S}_t$  is analytic.

To see this, write  $\mathcal{S}_t$  as an intersection of three sets,

$$\begin{aligned} B_1 &:= \left\{ (\omega^t, h) \mid h \in K_t^{P^t}(\omega^t) \right\} \times \mathfrak{M}_1(\Omega_1), \\ B_2 &:= \left\{ (\omega^t, P) \mid P \in \mathfrak{P}_t(\omega^t) \right\} \times \mathbb{R}^d, \\ B_3 &:= \left\{ (\omega^t, h, P) \mid E^P [\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, 0, \dots, 0, -h)] < 0 \right\}, \end{aligned}$$

and show that each of those is analytic.

The set  $B_1$  is Borel, as it is just a product of  $\mathfrak{M}_1(\Omega_1)$  and the graph of  $K_t^{P^t}(\omega^t)$ , which is Borel; see [24, Theorem 14.8, p. 648].

The set  $B_2$  is analytic being the product of  $\mathbb{R}^d$  and the graph of  $\mathfrak{P}_t$ , which is analytic by assumption.

To show that  $B_3$  is analytic, use the assumption that  $\Psi_{t+1}$  is a lower semianalytic map. By [3, Lemma 7.30(2), p. 177] also the map  $\Psi_{t+1}^\infty$  is; it is defined as a limit of lower semianalytic functions. By the same argument as in Lemma 30, we see that the set

$$\left\{ (\omega^t, h_0, \dots, h_t, P) \in \Omega^t \times \mathbb{R}^{d(t+1)} \times \mathfrak{M}_1(\Omega_1) \mid E^P [\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, -h_0, \dots, -h_t)] < 0 \right\}$$

is analytic. Denote the above set by  $\tilde{B}_3$ . The projection

$$\begin{aligned} \Pi: \Omega^{t+1} \times \mathbb{R}^{d(t+1)} \times \mathfrak{M}_1(\Omega_1) &\rightarrow \Omega^t \times \mathbb{R}^d \times \mathfrak{M}_1(\Omega_1) \\ (\omega^t, x_0, \dots, x_t, P) &\mapsto (\omega^t, x_t, P) \end{aligned}$$

is continuous and thus Borel. We deduce from [3, Proposition 7.40, p. 165] that the set

$$B_3 = \left\{ (\omega^t, h, P) \in \Omega^t \times \mathbb{R}^d \times \mathfrak{M}_1(\Omega_1) \mid E^P [\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, 0, \dots, 0, -h)] < 0 \right\}$$

is analytic, as

$$B_3 = \Pi \left( \tilde{B}_3 \cap (\Omega^t \times (\{0\}^{dt} \times \mathbb{R}^d) \times \mathfrak{M}_1(\Omega_1)) \right).$$

*Step 3.* The desired strategy can be obtained from the selection of  $\mathcal{S}_t$ .

Define the set

$$\text{proj } \mathcal{S}_t = \left\{ \omega^t \mid \mathcal{S}_t \cap (\{\omega^t\} \times \mathbb{R}^d \times \mathfrak{M}_1(\Omega_1)) \neq \emptyset \right\},$$

which is just the projection of the set  $\mathcal{S}_t$  onto the first coordinate. Let us first show that the sets  $\text{proj } \mathcal{S}_t$  and  $G_t^c$  are equal up to a  $P^t$  nullset. Recall that the probability measure  $P^t$  was chosen at the beginning of the proof. Then, by definition of the sets  $G_t^c$  and  $\mathcal{S}_t$  and as  $K_t^{P^t} = K_t$   $P^t$ -a.s., we have  $P^t[\text{proj } \mathcal{S}_t] = P^t[G_t^c] > 0$ .

From Step 2, we know that  $\mathcal{S}_t$  is analytic. Therefore, the Jankov-von Neumann theorem [3, Proposition 7.49, p. 182] implies the existence of a universally measurable map  $\omega^t \mapsto (h_t(\omega^t), P_t(\omega^t))$  such that  $(\omega^t, h_t(\omega^t), P_t(\omega^t)) \in \mathcal{S}_t$  for all  $\omega^t \in \text{proj } \mathcal{S}_t$ . On the universally measurable set  $\{h_t \notin K_t\} \subseteq \Omega^t$ , we set  $h_t := 0 \in \mathbb{R}^d$  to guarantee that  $h_t(\omega^t) \in K_t(\omega^t)$  for every  $\omega^t \in \Omega^t$ . In the same way, we can define  $P_t(\cdot)$  to be any measurable selector of  $\mathfrak{P}_t(\cdot)$  on  $\{h_t \notin K_t\}$ . Recall that  $\{h_t \notin K_t\} \cap \text{proj } \mathcal{S}_t$  is a  $P^t$ -nullset.

Finally, we claim for the strategy  $H^{t+1} := (0, \dots, 0, h_t) \in \mathcal{H}^{t+1}$  that  $H^{t+1} \in \mathcal{K}^{t+1}$ , but  $-H^{t+1} \notin \mathcal{K}^{t+1}$ . To see this, observe first that as  $h_t(\omega^t) \in K_t(\omega^t)$  for all  $\omega^t$ , we have by definition

$$\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, H^{t+1}(\omega^t)) \geq 0 \quad \mathfrak{P}_t(\omega^t)\text{-q.s.}$$

for all  $\omega^t$ . For every  $\bar{P} \in \mathfrak{P}$  denote its restriction to  $\Omega^t$  by  $\bar{P}^t$ , i.e.,  $\bar{P}^t := \bar{P}|_{\Omega^t}$ . By the definition of the set  $\mathfrak{P}$  we have  $\bar{P}^{t+1} = \bar{P}^t \otimes \bar{P}_t$  for some selector  $\bar{P}_t \in \mathfrak{P}_t$ . By Fubini's theorem we get that every  $\bar{P} \in \mathfrak{P}$  satisfies

$$\bar{P}[\Psi_{t+1}^\infty(H^{t+1}) \geq 0] = E^{\bar{P}^t(d\omega^t)}[\bar{P}_t(\omega^t)[\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, H^{t+1}(\omega^t)) \geq 0]] = 1,$$

which proves that  $H^{t+1} \in \mathcal{K}^{t+1}$ . To see that  $-H^{t+1} \notin \mathcal{K}^{t+1}$ , define the measure  $\tilde{P} \in \mathfrak{P}$  by

$$\tilde{P} := P^t \otimes P_t \otimes \tilde{P}_{t+1} \otimes \dots \otimes \tilde{P}_{T-1},$$

where  $P^t$  is the measure introduced at the beginning of the proof, the kernel  $P_t$  is the one selected from  $\mathcal{S}_t$  above, and  $\tilde{P}_s \in \mathfrak{P}_s$  are any selections of  $\mathfrak{P}_s$  for  $s := t+1, \dots, T-1$ . Recalling that  $\Psi_{t+1}^\infty \leq 0$ , by definition of  $h$  we get

$$E^{\tilde{P}}[\Psi_{t+1}^\infty(-H^t)] = E^{P^t(d\omega^t)}[E^{P_t(\omega^t)}[\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, -H^t(\omega^t))]] < 0,$$

as  $\tilde{P}[\text{proj } \mathcal{S}_t] > 0$ . Hence  $-H^{t+1} \notin \mathcal{K}^{t+1}$ , which gives us a contradiction to the linearity of  $\mathcal{K}^{t+1}$ .  $\square$

Now we will prove Proposition 27, which is, basically, a (measurable) version of Theorem 18 stating the existence of a (local) maximizer in the one-period model at time  $t$ .

*Proof of Proposition 27.* Recall that, by Remark 23, the map  $\tilde{\Psi}_t(\omega^t, \cdot)$  is concave, proper with  $h^{\circ, t} \in \mathbb{R}^{dt}$  in the interior of its domain for each  $\omega^t$ . Hence by [24, Theorem 2.35, p. 59], we have for each  $\omega^t$  that

$$\Psi_t(\omega^t, x^t) = \lim_{\lambda \nearrow 1} \tilde{\Psi}_t(\omega^t, \lambda x^t + (1 - \lambda)h^{\circ, t}).$$

To prove (i), we want to show that for  $\mathfrak{P}$ -quasi every  $\omega^t$  the mapping  $x^{t+1} \mapsto \Phi_t(\omega^t, x^{t+1})$  satisfies the conditions of Proposition 19 with the linear mapping  $A$  being just the restriction  $A(x^t, x_t) = x^t$ . Fix an  $\omega^t \in \Omega^t$ . We deduce from the identity in (11) that  $\Phi_t^\infty(\omega^t, 0, \dots, 0, x_t) \geq 0$  if and only if  $\Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, 0, x_t) \geq 0$   $\mathfrak{P}_t(\omega^t)$ -q.s., i.e., if  $x_t \in K_t(\omega^t)$ . We know from Proposition 26 that  $\text{NA}_t$  holds, which means that  $K_t(\omega^t)$  is linear for  $\mathfrak{P}$ -quasi every  $\omega^t$ . Thus, for  $\mathfrak{P}$ -quasi every  $\omega^t$ , the conditions of Proposition 19 are indeed satisfied and hence  $\tilde{\Psi}_t(\omega^t, \cdot)$  is  $\mathfrak{P}$ -q.s. an upper-semicontinuous function. Moreover, from the definition of  $\Psi_t(\omega^t, \cdot)$  being the upper-semicontinuous hull of  $\tilde{\Psi}_t(\omega^t, \cdot)$ , statement (i) follows.

We now prove (ii). By Lemma 30 and Remark 3 we know that  $\Phi_t$  is an  $\mathcal{F}_t$ -normal integrand. Having chosen a strategy  $H^t \in \mathcal{H}^t$ , [24, Proposition 14.45, p. 669] yields that the mapping  $\Phi^{H^t}(\omega^t, x) := \Phi_t(\omega^t, H^t(\omega^t), x)$  is an  $\mathcal{F}_t$ -normal integrand, too. Therefore, we deduce from [24, Theorem 14.37, p. 664] that the set-valued mapping  $\Upsilon: \Omega^t \rightrightarrows \mathbb{R}^d$  defined by

$$\Upsilon(\omega^t) := \operatorname{argmax} \Phi^{H^t}(\omega^t, \cdot)$$

admits an  $\mathcal{F}_t$ -measurable selector  $\widehat{h}_t$  on the universally measurable set  $\{\Upsilon \neq \emptyset\}$ . Extend  $\widehat{h}_t$  by setting  $\widehat{h}_t = 0$  on  $\{\Upsilon = \emptyset\}$ . As we know from Proposition 26 that  $\text{NA}_t$  holds, the attainment of the supremum in Proposition 19 gives that  $\{\Upsilon = \emptyset\}$  is a  $\mathfrak{P}$ -polar set. Thus, the result follows, as  $\Psi_t = \widetilde{\Psi}_t$   $\mathfrak{P}$ -q.s.  $\square$

Next, to see that  $\Psi_{t+1}$  satisfying Assumption 1 implies that  $\Psi_t$  does, too, it remains to show that  $\Psi_t$  is lower semianalytic. To that end, we first need the following useful lemma.

**LEMMA 32.** *Let  $g: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$  be a concave upper-semicontinuous function having  $(x^\circ, y^\circ) \in \mathbb{R}^n \times \mathbb{R}^m$  in the interior of its domain. Then, the function*

$$h(x) := \lim_{\lambda \nearrow 1} \sup_{y \in \mathbb{R}^m} g(\lambda x + (1-\lambda)x^\circ, y)$$

*is upper-semicontinuous and satisfies*

$$h(x) = \lim_{\lambda \nearrow 1} \sup_{y \in \mathbb{Q}^m} g(\lambda x + (1-\lambda)x^\circ, y).$$

*Proof.* Define the function

$$\widetilde{h}(x) := \sup_{y \in \mathbb{R}^m} g(x, y).$$

Then, by [24, Theorem 2.35, p. 59],  $h(x) = \text{cl} \widetilde{h}(x)$ , i.e.,  $h$  is the upper-semicontinuous hull of  $\widetilde{h}$ ; in particular, it is upper-semicontinuous. Now, denote by

$$\begin{aligned} \text{dom } g &:= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid g(x, y) > -\infty\}, \\ \text{dom } \widetilde{h} &:= \{x \in \mathbb{R}^n \mid \widetilde{h}(x) > -\infty\} \end{aligned}$$

the domains of the functions  $g$  and  $\widetilde{h}$ , respectively. We have to differentiate two cases.

*Case 1.* Let  $x \in \text{dom } \widetilde{h}$ . Then for each  $\lambda \in (0, 1)$  we have  $\widetilde{h}(\lambda x + (1-\lambda)x^\circ) > -\infty$  and also

$$D_{\lambda x + (1-\lambda)x^\circ} := \text{dom } g \cap \{\lambda x + (1-\lambda)x^\circ\} \times \mathbb{R}^m \neq \emptyset.$$

Denote by  $\Pi: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  the projection on the second component. The set  $\Pi D_x$  does not necessarily have a nonempty interior, but  $\Pi D_{\lambda x + (1-\lambda)x^\circ}$  has, by assumption that an open ball around  $(x^\circ, y^\circ)$  is included in the domain of  $g$ . Hence using [24, Theorem 2.35, p. 59] yields

$$h(\lambda x + (1-\lambda)x^\circ) = \widetilde{h}(\lambda x + (1-\lambda)x^\circ) = \sup_{y \in \mathbb{Q}^m} g(\lambda x + (1-\lambda)x^\circ, y).$$

Taking the limit as  $\lambda \nearrow 1$ , using the upper-semicontinuity of  $h$  proves the claim in the first case.

*Case 2.* Let  $x \notin \overline{\text{dom } \tilde{h}}$ . In this case, there exists a  $\lambda_m \in (0, 1)$ , such that  $\tilde{h}(\lambda x + (1 - \lambda)x^\circ) = -\infty$  for all  $\lambda > \lambda_m$ . This implies that the set  $D_{\lambda x + (1 - \lambda)x^\circ}$  defined above is empty for each  $\lambda > \lambda_m$ , which yields the claim.  $\square$

LEMMA 33. Fix  $t \in \{0, \dots, T - 1\}$ . If  $\Psi_{t+1}$  satisfies Assumption 1, then the map  $\Psi_t$  is lower semianalytic.

*Proof.* By Lemma 30, the map

$$\Omega^t \times \mathbb{R}^{dt} \times \mathbb{R}^d \rightarrow \overline{\mathbb{R}}, \quad (\omega^t, x^t, \tilde{x}) \mapsto \Phi_t(\omega^t, x^t, \tilde{x})$$

is lower semianalytic. Lemma 32 now yields

$$\Psi_t(\omega^t, x^t) = \lim_{\lambda \nearrow 1} \sup_{\tilde{x} \in \mathbb{Q}^d} \Phi_t(\omega^t, \lambda x^t + (1 - \lambda)h^{\circ, t}, \tilde{x}).$$

This implies that  $\Psi_t$  is lower semianalytic due to the fact that the countable supremum of lower semianalytic functions is again lower semianalytic and a limit of a sequence of lower semianalytic functions is again lower semianalytic; see [3, Lemma 7.30(2), p. 178].  $\square$

*Proof of Proposition 28.* The structure of the proof is similar to that one of Proposition 26. Assume by contradiction that  $\mathcal{K}^t$  is not linear. Then, there exists a probability measure in  $\mathfrak{P}$  with its restriction to  $\Omega^t$  denoted by  $P^t$ , and  $\tilde{H}^t \in \mathcal{H}^t$  such that

$$\Psi_t^\infty(\tilde{H}^t) \geq 0 \quad \mathfrak{P}\text{-q.s.} \quad \text{and} \quad P^t[\Psi_t^\infty(-\tilde{H}^t) < 0] > 0.$$

*Step 1.* We claim that there exists an  $\mathcal{F}_t$ -measurable map  $\tilde{h}_t : \Omega^t \rightarrow \mathbb{R}^d$  such that

$$\Phi_t^\infty(\tilde{H}^t, \tilde{h}_t) = \Psi_t^\infty(\tilde{H}^t) \quad \mathfrak{P}\text{-q.s.}$$

Indeed, applying Proposition 19 to the function  $f(\cdot) = \Phi_t^\infty(\omega^t, \cdot)$  yields that the set-valued map

$$M(\omega^t) := \left\{ h \in \mathbb{R}^d \mid \Phi_t^\infty(\omega^t, \tilde{H}^t, h) = \Psi_t^\infty(\omega^t, \tilde{H}^t) \right\}$$

is not empty for  $\mathfrak{P}$ -quasi every  $\omega^t$ . Then, following the proof of Proposition 27(ii) using [24, Theorem 14.37, p. 664] provides existence of an  $\mathcal{F}_t$ -measurable selector  $\tilde{h}_t$  of  $M$ .

*Step 2.* Let us show that  $\tilde{H}^{t+1} := (\tilde{H}^t, \tilde{h}_t) \in \mathcal{H}^{t+1}$  satisfies  $\Psi_{t+1}^\infty(\tilde{H}^{t+1}) \geq 0$   $\mathfrak{P}$ -q.s., i.e.,  $\tilde{H}^{t+1} \in \mathcal{K}^{t+1}$ .

For  $\mathfrak{P}$ -q.e.  $\omega^t$  we have that

$$\begin{aligned} 0 &= \Psi_t^\infty(\omega^t, \tilde{H}^t(\omega^t)) \\ &= \Phi_t^\infty(\omega^t, \tilde{H}^{t+1}(\omega^t)) = \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P \left[ \Psi_{t+1}^\infty(\omega^t \otimes_t \cdot, \tilde{H}^{t+1}(\omega^t)) \right]. \end{aligned}$$

As every  $P' \in \mathfrak{P}$  satisfies  $P'|_{\Omega^{t+1}} = P'|_{\Omega^t} \otimes P'_t$  for some selection  $P'_t \in \mathfrak{P}_t$ , we obtain the result directly from Fubini's theorem.

*Step 3.* We want to show that  $-\tilde{H}^{t+1} \notin \mathcal{K}^{t+1}$ . To see this, recall the probability measure  $P^t$  on  $\Omega^t$  introduced at the beginning of the proof. We first modify  $\tilde{H}^{t+1}$  on

a  $P^t$ -nullset to obtain a Borel measurable function  $\tilde{H}^{P^t, t+1}$ . Consider the set

$$\mathfrak{S}_t := \left\{ (\omega^t, P) \in \Omega^t \times \mathfrak{M}_1(\Omega_1) \mid P \in \mathfrak{P}_t(\omega^t), \right. \\ \left. E^P \left[ \Psi_{t+1}^\infty \left( \omega^t \otimes_t \cdot, -\tilde{H}^{P^t, t+1}(\omega^t) \right) \right] < 0 \right\}.$$

Using the same arguments as in Steps 2 and 3 of the proof of Proposition 26 we get that  $\mathfrak{S}_t$  is analytic, hence there exists a universally measurable kernel  $P_t : \Omega^t \mapsto \mathfrak{M}_1(\Omega_1)$  such that  $(\omega^t, P_t(\omega^t)) \in \mathfrak{S}_t$  for all  $\omega^t \in \text{proj } \mathfrak{S}_t$ .

We claim that  $P^t[\text{proj } \mathfrak{S}_t] > 0$ . To see this, observe first that  $\mathfrak{P}$ -q.s., we have for any  $x^t \in \mathbb{R}^{dt}$ ,  $x' \in \mathbb{R}^d$  that  $\Psi_t(x^t) \geq \Phi_t(x^t, x')$ . Hence, we obtain from [24, Theorem 3.21, p. 88] that  $\mathfrak{P}$ -q.s., we also have  $\Psi_t^\infty(x^t) \geq \Phi_t^\infty(x^t, x')$  for any  $x^t \in \mathbb{R}^{dt}$ ,  $x' \in \mathbb{R}^d$ . Therefore, we have for  $\mathfrak{P}$ -quasi every  $\omega^t$  that

$$\Psi_t^\infty \left( \omega^t, -\tilde{H}^t(\omega^t) \right) \geq \Phi_t^\infty \left( \omega^t, -\tilde{H}^{t+1}(\omega^t) \right) \\ = \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P \left[ \Psi_{t+1}^\infty \left( \omega^t \otimes_t \cdot, -\tilde{H}^{t+1}(\omega^t) \right) \right].$$

As  $P^t[\Psi_t^\infty(-\tilde{H}^t)] < 0] > 0$  and  $H^{t+1} = H^{P^t, t+1}$   $P^t$ -a.s., we conclude that indeed  $P^t[\text{proj } \mathfrak{S}_t] > 0$ . Finally, to see that  $-H^{t+1} \notin \mathcal{K}^{t+1}$ , define the measure  $\tilde{P} \in \mathfrak{P}$  by

$$\tilde{P} := P^t|_{\Omega^t} \otimes P_t \otimes \tilde{P}_{t+1} \otimes \cdots \otimes \tilde{P}_{T-1},$$

where we take any selector  $\tilde{P}_s \in \mathfrak{P}_s$  for  $s := t+1, \dots, T-1$ . Then, by construction, we have

$$E^{\tilde{P}} \left[ \Psi_{t+1}^\infty \left( -\tilde{H}^{t+1} \right) \right] < 0,$$

hence indeed  $-H^{t+1} \notin \mathcal{K}^{t+1}$ , which gives a contradiction to the linearity of  $\mathcal{K}^{t+1}$ .  $\square$

*Proof of Proposition 29.* We have shown that if  $\Psi_{t+1}$  satisfies Assumption 1 and  $\mathcal{K}^{t+1}$  is linear (i.e.,  $\text{NA}(\mathfrak{P})^{t+1}$  holds), then  $\Psi_t$  satisfies Assumption 1 and the local no-arbitrage condition  $\text{NA}_t$  holds. Moreover, we have shown that the linearity of  $\mathcal{K}^{t+1}$  implies the linearity of  $\mathcal{K}^t$  as soon as  $\Psi_{t+1}$  satisfies Assumption 1. As by assumption,  $\Psi = \Psi_T$  satisfies Assumption 1 and  $\text{NA}(\mathfrak{P})$  holds, we see that Proposition 29 holds by using backward induction.  $\square$

**5.2. Proof of Theorem 7.** The goal of this subsection is to give the proof of Theorem 7, which is the main result of this paper. We will construct the optimal strategy  $\hat{H} := (\hat{H}_0, \dots, \hat{H}_{T-1}) \in \mathcal{H}$  recursively from time  $t = 0$  upward by applying Proposition 27 at each time  $t$ , given the restricted strategy  $\hat{H}^t := (\hat{H}_0, \dots, \hat{H}_{t-1}) \in \mathcal{H}^t$ . We follow [18] to check that  $\hat{H}$  is indeed an optimizer of (6).

LEMMA 34. Let  $t \in \{0, \dots, T-1\}$  and  $H^{t+1} \in \mathcal{H}^{t+1}$ . Define the random variable

$$X(\omega^t) := \Phi_t(\omega^t, H^{t+1}(\omega^t)).$$

For any  $\varepsilon > 0$ , there exists a universally measurable kernel  $P_t^\varepsilon : \Omega^t \rightarrow \mathfrak{M}_1(\Omega_1)$  such that  $P_t^\varepsilon(\omega^t) \in \mathfrak{P}_t(\omega^t)$  for all  $\omega^t \in \Omega^t$  and

$$E^{P_t^\varepsilon(\omega^t)} \left[ \Psi_{t+1} \left( \omega^t \otimes_t \cdot, H^{t+1}(\omega^t) \right) \right] \leq \begin{cases} X(\omega^t) + \varepsilon & \text{if } X(\omega^t) > -\infty, \\ -\varepsilon^{-1} & \text{otherwise.} \end{cases}$$

*Proof.* For any  $x^{t+1} \in \mathbb{R}^{d(t+1)}$ , define the random variable

$$\Phi^{x^{t+1}}(\omega^t) := \Phi_t(\omega^t, x^{t+1}).$$

By the proof of Lemma 30, the map  $(\omega^t, P, x^{t+1}) \mapsto E^P[\Psi_{t+1}(\omega^t \otimes_t \cdot, x^{t+1})]$  is lower semianalytic. Moreover, by assumption,  $\text{graph}(\mathfrak{P}_t)$  is analytic. Therefore, by [3, Theorem 7.50, p. 184], it admits a universally measurable kernel  $(\omega^t, x^{t+1}) \mapsto \tilde{P}_t^\varepsilon(\omega^t, x^{t+1}) \in \mathfrak{P}_t(\omega^t)$  satisfying

$$E^{\tilde{P}_t^\varepsilon(\omega^t, x^{t+1})}[\Psi_{t+1}(\omega^t \otimes_t \cdot, x^{t+1})] \leq \begin{cases} \Phi^{x^{t+1}}(\omega^t) + \varepsilon & \text{if } \Phi_t^{x^{t+1}}(\omega^t) > -\infty, \\ -\varepsilon^{-1} & \text{otherwise.} \end{cases}$$

Setting  $P_t^\varepsilon(\omega^t) := \tilde{P}_t^\varepsilon(\omega^t, H^{t+1}(\omega^t))$  yields the result, as  $X(\omega^t) = \Phi^{H^{t+1}(\omega^t)}(\omega^t)$ .  $\square$

Now we are able to prove Theorem 7.

*Proof of Theorem 7.* By Theorem 18, there exists  $\hat{H}_0 \in \mathbb{R}^d$  such that

$$\inf_{P \in \mathfrak{P}_0} E^P[\Psi_1(\hat{H}_0)] = \sup_{x \in \mathbb{R}^d} \inf_{P \in \mathfrak{P}_0} E^P[\Psi_1(x)].$$

By a recursive application of Proposition 27, we can define an  $\mathcal{F}_t$ -measurable random variable  $\hat{H}_t$  such that

$$\inf_{P \in \mathfrak{P}_t(\omega^t)} E^P[\Psi_{t+1}(\omega^t \otimes_t \cdot, \hat{H}^t(\omega^{t-1}), \hat{H}_t(\omega^t))] = \Psi_t(\omega^t, \hat{H}^t(\omega^{t-1}))$$

for  $\mathfrak{P}$ -quasi-every  $\omega^t \in \Omega^t$ , for all  $t = 1, \dots, T-1$ . We claim that  $\hat{H} \in \mathcal{H}$  is optimal, i.e., satisfies (6). We first show that

$$(13) \quad \inf_{P \in \mathfrak{P}} E^P[\Psi_T(\hat{H})] \geq \Psi_0.$$

To that end, let  $t \in \{0, \dots, T-1\}$ . Let  $P \in \mathfrak{P}$ ; we write  $P = P_0 \otimes \dots \otimes P_{T-1}$  with kernels  $\mathfrak{P}_s : \Omega^s \rightarrow \mathfrak{M}_1(\Omega_1)$  satisfying  $P_s(\cdot) \in \mathfrak{P}_s(\cdot)$ . Therefore, by applying Fubini's theorem and the definition of  $\hat{H}$

$$\begin{aligned} & E^P[\Psi_{t+1}(\hat{H}_0, \dots, \hat{H}_t)] \\ &= E^{(P_0 \otimes \dots \otimes P_{t-1})(d\omega^t)}[E^{P_t(\omega^t)}[\Psi_{t+1}(\omega^t \otimes_t \cdot, \hat{H}^t(\omega^{t-1}), \hat{H}_t(\omega^t))]] \\ &\geq E^{(P_0 \otimes \dots \otimes P_{t-1})(d\omega^t)}\left[\inf_{P' \in \mathfrak{P}_t(\omega^t)} E^{P'}[\Psi_{t+1}(\omega^t \otimes_t \cdot, \hat{H}^t(\omega^{t-1}), \hat{H}_t(\omega^t))]\right] \\ &= E^{(P_0 \otimes \dots \otimes P_{t-1})}[\Psi_t(\hat{H}^t)] \\ &= E^P[\Psi_t(\hat{H}^t)]. \end{aligned}$$

Using this inequality repeatedly from  $t = T-1$  to  $t = 0$  yields  $E^P[\Psi_T(\hat{H})] \geq \Psi_0$ . As  $P \in \mathfrak{P}$  was arbitrarily chosen, the claim (13) is proven. It remains to show that

$$\Psi_0 \geq \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P[\Psi(H)]$$



to see that  $\widehat{H} \in \mathcal{H}$  is optimal. So, fix an arbitrary  $H \in \mathcal{H}$ . It suffices to show that for every  $t \in \{0, \dots, T-1\}$

$$(14) \quad \inf_{P \in \mathfrak{P}} E^P [\Psi_t (H^t)] \geq \inf_{P \in \mathfrak{P}} E^P [\Psi_{t+1} (H^{t+1})].$$

Indeed, using the inequality repeatedly from  $t = 0$  until  $t = T - 1$  yields

$$\Psi_0 \geq \inf_{P \in \mathfrak{P}} E^P [\Psi_T(H)].$$

Furthermore, as  $H \in \mathcal{H}$  was arbitrary and  $\Psi_T = \Psi$ , we obtain the desired inequality

$$\Psi_0 \geq \sup_{H \in \mathcal{H}} \inf_{P \in \mathfrak{P}} E^P [\Psi(H)].$$

Now, to prove the inequality in (14), fix an  $\varepsilon > 0$ . By Lemma 34, there exists a kernel  $P_t^\varepsilon : \Omega^t \rightarrow \mathfrak{M}_1(\Omega_1)$  such that for all  $\omega^t \in \Omega^t$

$$\begin{aligned} & E^{P_t^\varepsilon(\omega^t)} [\Psi_{t+1} (\omega^t \otimes_t \cdot, H^{t+1}(\omega^t))] - \varepsilon \\ & \leq (-\varepsilon^{-1}) \vee \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P [\Psi_{t+1} (\omega^t \otimes_t \cdot, H^{t+1}(\omega^t))] \\ & \leq (-\varepsilon^{-1}) \vee \sup_{x \in \mathbb{R}^d} \inf_{P \in \mathfrak{P}_t(\omega^t)} E^P [\Psi_{t+1} (\omega^t \otimes_t \cdot, H^t(\omega^{t-1}), x)] \\ & = (-\varepsilon^{-1}) \vee \Psi_t (\omega^t, H^t(\omega^{t-1})). \end{aligned}$$

Take any  $P \in \mathfrak{P}$  and denote its restriction to  $\Omega^t$  by  $P^t$ . Integrating the above inequalities yields

$$E^{P^t} [(-\varepsilon^{-1}) \vee \Psi_t (H^t)] \geq E^{P^t \otimes P_t^\varepsilon} [\Psi_{t+1} (H^{t+1})] - \varepsilon \geq \inf_{P' \in \mathfrak{P}} E^{P'} [\Psi_{t+1} (H^{t+1})] - \varepsilon.$$

Letting  $\varepsilon \rightarrow 0$ , we obtain, by Fatou's lemma, that

$$E^P [\Psi_t (H^t)] \geq \inf_{P' \in \mathfrak{P}} E^{P'} [\Psi_{t+1} (H^{t+1})].$$

This implies the inequality (14), as  $P \in \mathfrak{P}$  was arbitrary.  $\square$

**Appendix.** Here we provide a simple fact about horizon functions of compositions. Let us first recall the definition of the domain of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{dom } f := \{x \in \mathbb{R}^n \mid f(x) > -\infty\}.$$

LEMMA 35. *Let  $U : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$  be concave, nondecreasing, nonconstant, and upper-semicontinuous. Let  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  be concave and upper-semicontinuous and assume that  $V(\mathbb{R}^n) \cap \text{dom } U \neq \emptyset$ . Then the function*

$$\Psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}, \quad h \mapsto \Psi(h) := \begin{cases} U(V(h)) & \text{if } h \in \text{dom } V, \\ -\infty & \text{otherwise} \end{cases}$$

*is concave, proper, and upper-semicontinuous. Moreover,  $\Psi^\infty$  satisfies*

$$\Psi^\infty(h) = \begin{cases} U^\infty(V^\infty(h)) & \text{if } h \in \text{dom } V^\infty, \\ -\infty & \text{otherwise.} \end{cases}$$

*Proof.* The first part of the lemma is obvious. The only thing requiring a proof is the statement about the form of the horizon function  $\Psi^\infty$ .

First, choose a point  $x \in \mathbb{R}^n$  such that  $V(x) \in \text{dom } U$ , or equivalently, such that  $\Psi(x) > -\infty$ . Now, fix any  $h \in \mathbb{R}^n$ . The mapping  $V$  is concave, hence

the sequence  $m \mapsto \frac{1}{m}(V(x + mh) - V(x)) =: a_m$  is nonincreasing.

Denote its limit by  $a$ . Then,  $a > -\infty$  if and only if  $h \in \text{dom } V^\infty$ ; indeed  $a = V^\infty(h)$ .

Let us first estimate the horizon function  $\Psi^\infty$  from above. If  $a_m = -\infty$  for some  $m$ , this implies by definition that  $\Psi^\infty(h) = -\infty$ , as  $(a_m)$  is nonincreasing. Now assume that  $a_m > -\infty$  for each  $m$ . Then

$$\begin{aligned}\Psi^\infty(h) &= \lim_{n \rightarrow \infty} \frac{1}{n}(\Psi(x + nh) - \Psi(x)) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \left( U \left( n \frac{1}{n}(V(x + nh) - V(x)) + V(x) \right) - \Psi(x) \right) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} (U(n a_m + V(x)) - U(V(x))) \\ &= U^\infty(a_m).\end{aligned}$$

If  $h \in \text{dom } V^\infty$ , then by letting  $m$  tend to infinity, the above estimate and upper-semicontinuity of  $U^\infty$  yield  $\Psi^\infty(h) \leq U^\infty(V^\infty(h))$ . If  $h \notin \text{dom } V^\infty$ , then  $a_m$  diverges to  $-\infty$ , hence also  $U(a_m)$  tends to  $-\infty$  as  $U$  is concave, nondecreasing, and nonconstant. By definition of the horizon function we have

$$U^\infty(a_m) \leq U(a_m + V(x)) - U(V(x)),$$

hence also  $U^\infty(a_m)$  tends to  $-\infty$ . This proves the desired first inequality.

To estimate  $\Psi^\infty$  from below, we only need to consider  $h \in \text{dom } V^\infty$ . Indeed, we know from above that  $\Psi^\infty(h) = -\infty$  whenever  $h \notin \text{dom } V^\infty$ . So let  $h \in \text{dom } V^\infty$ . Then

$$\begin{aligned}\Psi^\infty(h) &= \lim_{n \rightarrow \infty} \frac{1}{n} (U(n a_n + V(x)) - U(V(x))) \geq \lim_{n \rightarrow \infty} \frac{1}{n} (U(n a + V(x)) - U(V(x))) \\ &= U^\infty(V^\infty(h)).\end{aligned}\quad \square$$

## REFERENCES

- [1] R. ALMGREN AND N. CHRISS, *Optimal execution of portfolio transactions*, J. Risk, 3 (2000), pp. 5–39.
- [2] D. BARTL, *Exponential Utility Maximization Under Model Uncertainty for Unbounded Endowments*, preprint, arXiv:1610.00999, 2016.
- [3] D. P. BERTSEKAS AND S. E. SHREVE, *Stochastic Optimal Control. The Discrete-Time Case*, Academic Press, New York, 1978.
- [4] S. BIAGINI AND M. PINAR, *The robust Merton problem of an ambiguity averse investor*, Math. Financ. Econ., 11 (2017), 1–24.
- [5] B. BOUCHARD, *No-arbitrage in discrete-time markets with proportional transaction costs and general information structure*, Finance Stoch., 10 (2006), pp. 276–297.
- [6] B. BOUCHARD AND M. NUTZ, *Arbitrage and duality in nondominated discrete-time models*, Ann. Appl. Probab., 25 (2015), pp. 823–859.
- [7] L. CARASSUS AND R. BLANCHARD, *Robust Optimal Investment in Discrete Time for Unbounded Utility Function*, preprint, arXiv:1609.09205, 2016.
- [8] F. DELBAEN AND W. SCHACHERMAYER, *The Mathematics of Arbitrage*, Springer, Berlin, 2006.

- [9] L. DENIS AND M. KERVAREC, *Optimal investment under model uncertainty in nondominated models*, SIAM J. Control Optim., 51 (2013), pp. 1803–1822.
- [10] Y. DOLINSKY AND H. M. SONER, *Robust hedging with proportional transaction costs*, Finance Stoch., 18 (2014), pp. 327–347.
- [11] I. V. EVSTIGNEEV, *Measurable selection and dynamic programming*, Math. Oper. Res., 1 (1976), pp. 267–272.
- [12] J.-P. FOUQUE, C. S. PUN, AND H. Y. WONG, *Portfolio optimization with ambiguous correlation and stochastic volatilities*, SIAM J. Control Optim., 54 (2016), pp. 2309–2338.
- [13] Y. KABANOV, *Hedging and liquidation under transaction costs in currency markets*, Finance Stoch., 3 (1999), pp. 237–248.
- [14] Y. KABANOV AND C. STRICKER, *The Dalang-Morton-Willinger theorem under delayed and restricted information*, in In Memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX, Lecture Notes in Math. 1874, Springer, Berlin, 2006, pp. 209–213.
- [15] Q. LIN AND F. RIEDEL, *Optimal Consumption and Portfolio Choice with Ambiguity*, preprint arXiv:1401.1639v1, 2014.
- [16] A. MATOUSSI, D. POSSAMAI, AND C. ZHOU, *Robust utility maximization in non-dominated models with 2BSDEs: The uncertain volatility model*, Math. Finance, 25 (2015), pp. 258–287.
- [17] A. NEUFELD AND M. NUTZ, *Robust utility maximization with Lévy processes*, Math. Finance, 28 (2018), pp. 82–105.
- [18] M. NUTZ, *Utility maximization under model uncertainty in discrete time*, Math. Finance, 26 (2016), pp. 252–268.
- [19] K. OWARI, *A robust version of convex integral functionals*, J. Convex Anal., 22 (2015), pp. 827–852.
- [20] T. PENNANEN, *Convex duality in stochastic optimization and mathematical finance*, Math. Oper. Res., 36 (2011), pp. 340–362.
- [21] T. PENNANEN AND A.-P. PERKKIÖ, *Stochastic programs without duality gaps*, Math. Program., 136 (2012), pp. 91–110.
- [22] M. RÁSONYI AND L. STETTNER, *On the existence of optimal portfolios for the utility maximization problem in discrete time financial market models*, in From Stochastic Calculus to Mathematical Finance, Springer, Berlin, 2006, pp. 589–608.
- [23] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997.
- [24] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer, Berlin, 2009.
- [25] W. SCHACHERMAYER, *The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time*, Math. Finance, 14 (2004), pp. 19–48.
- [26] R. TEVZADZE, T. TORONJADZE, AND T. UZUNASHVILI, *Robust utility maximization for a diffusion market model with misspecified coefficients*, Finance. Stoch., 17 (2013), pp. 535–563.